

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model
Run on: June 8, 2004, 16:25:55 ; Search time 44 Seconds
(without alignments)
1163.719 Million cell updates/sec

Title: US-10-054-988-114
Perfect score: 962
Sequence: 1 MEPQLGPEAAALRPGWLALL.....DLVQDCHQGRKFLCMLR 182

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1155919 seqs, 281338677 residues

Total number of hits satisfying chosen parameters: 1155919

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA: *
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep: *
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep: *
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep: *
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep: *
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep: *
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep: *
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep: *
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep: *
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep: *
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep: *
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep: *
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep: *
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep: *
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep: *
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep: *
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep: *
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep: *
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	962	100.0	182	9	US-09-739-254-114 Sequence 114, App
2	962	100.0	182	9	US-09-904-615-114 Sequence 114, App
3	962	100.0	182	12	US-10-211-462-217 Sequence 217, App
4	962	100.0	182	14	US-10-054-988-114 Sequence 114, App
5	962	100.0	182	14	US-10-177-293-150 Sequence 150, App
6	962	100.0	182	14	US-10-055-098-114 Sequence 114, App
7	962	100.0	209	9	US-09-739-254-168 Sequence 168, App
8	962	100.0	209	9	US-09-904-615-168 Sequence 168, App
9	962	100.0	209	14	US-10-054-988-168 Sequence 168, App
10	962	100.0	209	14	US-10-055-098-168 Sequence 168, App
11	959	99.7	182	9	US-09-965-528-10 Sequence 10, Appl
12	959	99.7	182	12	US-10-147-493-172 Sequence 172, App
13	959	99.7	182	12	US-10-145-127-172 Sequence 172, App
14	959	99.7	182	12	US-10-160-503-172 Sequence 172, App
15	959	99.7	182	12	US-10-143-118-172 Sequence 172, App

16	959	99.7	182	12	US-10-144-993-172	Sequence 172, App
17	959	99.7	182	12	US-10-158-787-172	Sequence 172, App
18	959	99.7	182	12	US-09-969-984-10	Sequence 10, Appl
19	959	99.7	182	12	US-10-140-024-172	Sequence 172, App
20	959	99.7	182	12	US-10-140-808-172	Sequence 172, App
21	959	99.7	182	12	US-10-152-405-172	Sequence 172, App
22	959	99.7	182	12	US-10-127-852A-172	Sequence 172, App
23	959	99.7	182	12	US-10-127-900A-172	Sequence 172, App
24	959	99.7	182	12	US-10-128-685A-172	Sequence 172, App
25	959	99.7	182	12	US-10-131-820A-172	Sequence 172, App
26	959	99.7	182	12	US-10-142-886-172	Sequence 172, App
27	959	99.7	182	12	US-10-146-728-172	Sequence 172, App
28	959	99.7	182	12	US-10-146-786-172	Sequence 172, App
29	959	99.7	182	12	US-10-147-499-172	Sequence 172, App
30	959	99.7	182	12	US-10-157-798-172	Sequence 172, App
31	959	99.7	182	14	US-10-028-072-172	Sequence 172, App
32	959	99.7	182	14	US-10-121-049-172	Sequence 172, App
33	959	99.7	182	14	US-10-123-904-172	Sequence 172, App
34	959	99.7	182	14	US-10-140-470-172	Sequence 172, App
35	959	99.7	182	14	US-10-175-746-172	Sequence 172, App
36	959	99.7	182	14	US-10-176-918-172	Sequence 172, App
37	959	99.7	182	14	US-10-176-921-172	Sequence 172, App
38	959	99.7	182	14	US-10-137-865-172	Sequence 172, App
39	959	99.7	182	14	US-10-140-474-172	Sequence 172, App
40	959	99.7	182	14	US-10-142-431-172	Sequence 172, App
41	959	99.7	182	14	US-10-143-114-172	Sequence 172, App
42	959	99.7	182	14	US-10-140-002-172	Sequence 172, App
43	959	99.7	182	14	US-10-142-419-172	Sequence 172, App
44	959	99.7	182	14	US-10-123-262-172	Sequence 172, App
45	959	99.7	182	14	US-10-142-423-172	Sequence 172, App

ALIGNMENTS

RESULT 1
US-09-739-254-114
; Sequence 114, Application US/09739254
; Patent No. US20010021700A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 49 Human Secreted Proteins
; FILE REFERENCE: PZ032P1
; CURRENT APPLICATION NUMBER: US/09/739,254
; CURRENT FILING DATE: 2000-12-19
; EARLIER APPLICATION NUMBER: 09/511,554
; EARLIER FILING DATE: 2000-02-23
; EARLIER APPLICATION NUMBER: PCT/US99/19330
; EARLIER FILING DATE: 1999-08-24
; EARLIER APPLICATION NUMBER: 60/097,917
; EARLIER FILING DATE: 1998-08-25
; EARLIER APPLICATION NUMBER: 60/098,634
; EARLIER FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 170
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 114
; LENGTH: 182
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-739-254-114

Query Match 100.0%; Score 962; DB 9; Length 182;
Best Local Similarity 100.0%; Pred. No. 7.5e-97;
Matches 182; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MEPQLGPEAAALRPGWLALLWVSALSCSFSLPASSLSLVLPQVRTSYNFGRTFLGLDKC	60
Db	1	MEPQLGPEAAALRPGWLALLWVSALSCSFSLPASSLSLVLPQVRTSYNFGRTFLGLDKC	60
Qy	61	NACIGTSICKKFKKEEIRSDNWLASHLGLPPDLSLLSYNPANYSDDSKIWPRPVEIFRLVSKY	120
Db	61	NACIGTSICKKFKKEEIRSDNWLASHLGLPPDLSLLSYNPANYSDDSKIWPRPVEIFRLVSKY	120

QY 121 QNEISDRKICASASAPKTCISIERVLRKTERFQKWLOAKRLTPDLVQDCHQGQRELKFLCM 180
Db 121 QNEISDRKICASASAPKTCISIERVLRKTERFQKWLOAKRLTPDLVQDCHQGQRELKFLCM 180
QY 181 LR 182
Db 181 LR 182

RESULT 2

US-09-904-615-114
; Sequence 114, Application US/09904615
; Patent No. US2002026040A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 49 Human Secreted Proteins
; FILE REFERENCE: PZ032P1
; CURRENT APPLICATION NUMBER: US/09/904,615
; CURRENT FILING DATE: 2001-07-16
; PRIOR APPLICATION NUMBER: 09/511,554
; PRIOR FILING DATE: 2000-02-23
; PRIOR APPLICATION NUMBER: 60/097,917
; PRIOR FILING DATE: 1998-08-25
; PRIOR APPLICATION NUMBER: 60/098,634
; PRIOR FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 170
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 114
; LENGTH: 182
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-904-615-114

Query Match 100.0%; Score 962; DB 9; Length 182;
Best Local Similarity 100.0%; Pred. No. 7.5e-97;
Matches 182; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MEPQLGPEAAALRPGWLALLLWVSALSCSFSLPASSLSLVPQVRTSYNFGRTFLGLDKC 60
Db 1 MEPQLGPEAAALRPGWLALLLWVSALSCSFSLPASSLSLVPQVRTSYNFGRTFLGLDKC 60
QY 61 NACIGTSICKKFFKKEIRSDNWLASHLGLPPDLSLSPANYSDSDSKIWRPVEIFRLVSKY 120
Db 61 NACIGTSICKKFFKKEIRSDNWLASHLGLPPDLSLSPANYSDSDSKIWRPVEIFRLVSKY 120
QY 121 QNEISDRKICASASAPKTCISIERVLRKTERFQKWLOAKRLT2DLVQDCHQGQRELKFLCM 180
Db 121 QNEISDRKICASASAPKTCISIERVLRKTERFQKWLOAKRLTPDLVQDCHQGQRELKFLCM 180
QY 181 LR 182
Db 181 LR 182

RESULT 3

US-10-211-462-217
; Sequence 217, Application US/10211462
; Publication No. US20040033495A1
; GENERAL INFORMATION:
; APPLICANT: Murray, Richard
; APPLICANT: Glynn, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Aziz, Natasha
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Angiogenesis, Compositions and
; TITLE OF INVENTION: Methods of Screening for Angiogenesis Modulators
; FILE REFERENCE: 018501-006200US
; CURRENT APPLICATION NUMBER: US/10/211,462
; CURRENT FILING DATE: 2003-02-13
; PRIOR APPLICATION NUMBER: US 09/784,356
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: US 09/791,390
; PRIOR FILING DATE: 2001-02-22

; PRIOR APPLICATION NUMBER: US 60/310,025
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: US 60/334,244
; PRIOR FILING DATE: 2001-11-29
; NUMBER OF SEQ ID NOS: 230
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 217
; LENGTH: 182
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-211-462-217

Query Match 100.0%; Score 962; DB 12; Length 182;
Best Local Similarity 100.0%; Pred. No. 7.5e-97;
Matches 182; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MEPQLGPEAAALRPGWLALLLWVSALSCSFSLPASSLSLVPQVRTSYNFGRTFLGLDKC 60
Db 1 MEPQLGPEAAALRPGWLALLLWVSALSCSFSLPASSLSLVPQVRTSYNFGRTFLGLDKC 60
QY 61 NACIGTSICKKFFKKEIRSDNWLASHLGLPPDLSLSPANYSDSDSKIWRPVEIFRLVSKY 120
Db 61 NACIGTSICKKFFKKEIRSDNWLASHLGLPPDLSLSPANYSDSDSKIWRPVEIFRLVSKY 120
QY 121 QNEISDRKICASASAPKTCISIERVLRKTERFQKWLOAKRLTPDLVQDCHQGQRELKFLCM 180
Db 121 QNEISDRKICASASAPKTCISIERVLRKTERFQKWLOAKRLTPDLVQDCHQGQRELKFLCM 180
QY 181 LR 182
Db 181 LR 182

RESULT 4

US-10-054-988-114
; Sequence 114, Application US/10054988
; Publication No. US20030087341A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 49 Human Secreted Proteins
; FILE REFERENCE: PZ032P1
; CURRENT APPLICATION NUMBER: US/10/054,988
; CURRENT FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: 09/904,615
; PRIOR FILING DATE: 2001-07-16
; PRIOR APPLICATION NUMBER: 09/511,554
; PRIOR FILING DATE: 2000-02-23
; PRIOR APPLICATION NUMBER: 60/097,917
; PRIOR FILING DATE: 1998-08-25
; PRIOR APPLICATION NUMBER: 60/098,634
; PRIOR FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 170
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 114
; LENGTH: 182
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-054-988-114

Query Match 100.0%; Score 962; DB 14; Length 182;
Best Local Similarity 100.0%; Pred. No. 7.5e-97;
Matches 182; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MEPQLGPEAAALRPGWLALLLWVSALSCSFSLPASSLSLVPQVRTSYNFGRTFLGLDKC 60
Db 1 MEPQLGPEAAALRPGWLALLLWVSALSCSFSLPASSLSLVPQVRTSYNFGRTFLGLDKC 60
QY 61 NACIGTSICKKFFKKEIRSDNWLASHLGLPPDLSLSPANYSDSDSKIWRPVEIFRLVSKY 120
Db 61 NACIGTSICKKFFKKEIRSDNWLASHLGLPPDLSLSPANYSDSDSKIWRPVEIFRLVSKY 120
QY 121 QNEISDRKICASASAPKTCISIERVLRKTERFQKWLOAKRLTPDLVQDCHQGQRELKFLCM 180

Db 121 QNEISDRKICASASAPKTCISIERVLRKTERFQKWLOAKRLTPDLVQDCHQGORELKFLCM 180

QY 181 LR 182

||

Db 181 LR 182

RESULT 5

US-10-177-293-150
; Sequence 150, Application US/10177293
; Publication No. US20030124128A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Ghatt, Karen
; APPLICANT: Zhao, Xumei
; APPLICANT: Gannavarpu, Manjula
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Mertens, Maureen
; APPLICANT: Myer, Vic
; APPLICANT: Wang, Youzhen
; APPLICANT: Xu, Yongyao
; APPLICANT: Hoersch, Sebastian
; APPLICANT: Monahan, John
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Bast Jr., Robert C.
; APPLICANT: Hortobagyi, Gabriel N.
; APPLICANT: Pusztai, Lajos
; APPLICANT: Meric, Funda
; APPLICANT: Sahin, Aysegul
; APPLICANT: Mills, Gordon B.
; TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION, ASSESSMENT,
; TITLE OF INVENTION: PREVENTION, AND THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-038
; CURRENT APPLICATION NUMBER: US/10/177,293
; CURRENT FILING DATE: 2002-06-21
; PRIOR APPLICATION NUMBER: US 60/299,887
; PRIOR FILING DATE: 2001-06-21
; PRIOR APPLICATION NUMBER: US 60/301,572
; PRIOR FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: US 60/306,501
; PRIOR FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: US 60/325,002
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 60/362,585
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/xxx,xxx
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 506
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 150
; LENGTH: 182
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-177-293-150

Query Match 100.0%; Score 962; DB 14; Length 182;
Best Local Similarity 100.0%; Pred. No. 7.5e-97;
Matches 182; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEPOLGPEAAALRPGWLALLLWVSALSCSFLPASSLSLVPQVRTSYNFGRTFLGLDKC 60

|||||

Db 1 MEPOLGPEAAALRPGWLALLLWVSALSCSFLPASSLSLVPQVRTSYNFGRTFLGLDKC 60

QY 61 NACIGTSICKKFKKEIRSDNWLASHLGLPPDLSLSPANYSDDSKIWRPVEIFRLVSKY 120

|||||

Db 61 NACIGTSICKKFKKEIRSDNWLASHLGLPPDLSLSPANYSDDSKIWRPVEIFRLVSKY 120

QY 121 QNEISDRKICASAPKTCISIERVLRKTERFQKWLOAKRLTPDLVQDCHQGORELKFLCM 180

|||||

Db 121 QNEISDRKICASAPKTCISIERVLRKTERFQKWLOAKRLTPDLVQDCHQGORELKFLCM 180

QY 181 LR 182

||

Db 181 LR 182

RESULT 6

US-10-055-098-114
; Sequence 114, Application US/10055098
; Publication No. US20030139954A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 49 Human Secreted Proteins
; FILE REFERENCE: PZ032P1
; CURRENT APPLICATION NUMBER: US/10/055,098
; CURRENT FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/511,554
; PRIOR FILING DATE: EARLIER FILING DATE: 2000-02-23
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: PCT/US99/19330
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-08-24
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/097,917
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-08-25
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/098,634
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 170
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 114
; LENGTH: 182
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-055-098-114

Query Match 100.0%; Score 962; DB 14; Length 182;
Best Local Similarity 100.0%; Pred. No. 7.5e-97;
Matches 182; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEPOLGPEAAALRPGWLALLLWVSALSCSFLPASSLSLVPQVRTSYNFGRTFLGLDKC 60

|||||

Db 1 MEPOLGPEAAALRPGWLALLLWVSALSCSFLPASSLSLVPQVRTSYNFGRTFLGLDKC 60

QY 61 NACIGTSICKKFKKEIRSDNWLASHLGLPPDLSLSPANYSDDSKIWRPVEIFRLVSKY 120

|||||

Db 61 NACIGTSICKKFKKEIRSDNWLASHLGLPPDLSLSPANYSDDSKIWRPVEIFRLVSKY 120

QY 121 QNEISDRKICASAPKTCISIERVLRKTERFQKWLOAKRLTPDLVQDCHQGORELKFLCM 180

|||||

Db 121 QNEISDRKICASAPKTCISIERVLRKTERFQKWLOAKRLTPDLVQDCHQGORELKFLCM 180

QY 181 LR 182

||

Db 181 LR 182

RESULT 7

US-09-739-254-168
; Sequence 168, Application US/09739254
; Patent No. US20010321700A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 49 Human Secreted Proteins
; FILE REFERENCE: PZ032P1
; CURRENT APPLICATION NUMBER: US/09/739,254
; CURRENT FILING DATE: 2000-12-19
; EARLIER APPLICATION NUMBER: 09/511,554
; EARLIER FILING DATE: 2000-02-23
; EARLIER APPLICATION NUMBER: PCT/US99/19330
; EARLIER FILING DATE: 1999-08-24
; EARLIER APPLICATION NUMBER: 60/097,917
; EARLIER FILING DATE: 1998-08-25
; EARLIER APPLICATION NUMBER: 60/098,634
; EARLIER FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 170
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 168
; LENGTH: 209
; TYPE: PRT


```
; APPLICANT: Gurney,Austin L.
; APPLICANT: Sherwood,Steven
; APPLICANT: Smith,Victoria
; APPLICANT: Stewart,Timothy A.
; APPLICANT: Tumas,Daniel
; APPLICANT: Watanabe,Colin K
; APPLICANT: Wood,William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330R1C252
; CURRENT APPLICATION NUMBER: US/10/145,127
; CURRENT FILING DATE: 2002-05-13
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 172
; LENGTH: 182
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-145-127-172
```

```
Query Match 99.7%; Score 959; DB 12; Length 182;
Best Local Similarity 99.5%; Pred. No. 1.6e-96;
Matches 181; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEPQLGPEAAALRPGWLALLLWVSALSCSFSLPASSLSLVLPQVRTSYNFGRTFLGLDKC 60
Db 1 MEPQLGPEAAALRPGWLALLLWVSALSCSFSLPASSLSLVLPQVRTSYNFGRTFLGLDKC 60

QY 61 NACIGTSICKKFFKEEIRSDNWLASHLGLPPDLSLSPANYSDDSKIWRPVEIFRLVSKY 120
Db 61 NACIGTSICKKFFKEEIRSDNWLASHLGLPPDLSLSPANYSDDSKIWRPVEIFRLVSKY 120

QY 121 QNEISDRKICASAPKTCSTIERVLRKTERFQKWLQAKRLTPDLVQDCHQOGRELKFLCM 180
Db 121 QNEISDRRICASAPKTCSTIERVLRKTERFQKWLQAKRLTPDLVQDCHQOGRELKFLCM 180

QY 181 LR 182
Db 181 LR 182
```

```
RESULT 14
US-10-160-503-172
; Sequence 172, Application US/10160503
; Publication No. US20040033559A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C446
; CURRENT APPLICATION NUMBER: US/10/160,503
; CURRENT FILING DATE: 2002-05-30
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 172
; LENGTH: 182
; TYPE: PRT
```

```
; ORGANISM: Homo Sapien
US-10-160-503-172

Query Match 99.7%; Score 959; DB 12; Length 182;
Best Local Similarity 99.5%; Pred. No. 1.6e-96;
Matches 181; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEPQLGPEAAALRPGWLALLLWVSALSCSFSLPASSLSLVLPQVRTSYNFGRTFLGLDKC 60
Db 1 MEPQLGPEAAALRPGWLALLLWVSALSCSFSLPASSLSLVLPQVRTSYNFGRTFLGLDKC 60

QY 61 NACIGTSICKKFFKEEIRSDNWLASHLGLPPDLSLSPANYSDDSKIWRPVEIFRLVSKY 120
Db 61 NACIGTSICKKFFKEEIRSDNWLASHLGLPPDLSLSPANYSDDSKIWRPVEIFRLVSKY 120

QY 121 QNEISDRKICASAPKTCSTIERVLRKTERFQKWLQAKRLTPDLVQDCHQOGRELKFLCM 180
Db 121 QNEISDRRICASAPKTCSTIERVLRKTERFQKWLQAKRLTPDLVQDCHQOGRELKFLCM 180

QY 181 LR 182
Db 181 LR 182
```

```
RESULT 15
US-10-143-118-172
; Sequence 172, Application US/10143118
; Publication No. US20040038335A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C228
; CURRENT APPLICATION NUMBER: US/10/143,118
; CURRENT FILING DATE: 2002-05-09
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 172
; LENGTH: 182
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-143-118-172
```

```
Query Match 99.7%; Score 959; DB 12; Length 182;
Best Local Similarity 99.5%; Pred. No. 1.6e-96;
Matches 181; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEPQLGPEAAALRPGWLALLLWVSALSCSFSLPASSLSLVLPQVRTSYNFGRTFLGLDKC 60
Db 1 MEPQLGPEAAALRPGWLALLLWVSALSCSFSLPASSLSLVLPQVRTSYNFGRTFLGLDKC 60

QY 61 NACIGTSICKKFFKEEIRSDNWLASHLGLPPDLSLSPANYSDDSKIWRPVEIFRLVSKY 120
Db 61 NACIGTSICKKFFKEEIRSDNWLASHLGLPPDLSLSPANYSDDSKIWRPVEIFRLVSKY 120

QY 121 QNEISDRKICASAPKTCSTIERVLRKTERFQKWLQAKRLTPDLVQDCHQOGRELKFLCM 180
Db 121 QNEISDRRICASAPKTCSTIERVLRKTERFQKWLQAKRLTPDLVQDCHQOGRELKFLCM 180
```


Db 121 ONEISDRRICASAPKTCISIERVLRKTERFPQKWLQAKRLTPDLVQDCHQGORELKFLCM 180
Qy 181 LR 182
Db 181 LR 182

Search completed: June 8, 2004, 16:31:40
Job time : 45 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 8, 2004, 16:23:20 ; Search time 22 Seconds
(without alignments)
427.088 Million cell updates/sec

Title: US-10-054-988-114
Perfect score: 962
Sequence: 1 MEPQLGPEAAALPGWLALL.....DLVQDCHQGRELKFLCMLR 182

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	962	100.0	182	4	US-09-904-615-114 Sequence 114, App
2	962	100.0	209	4	US-09-904-615-168 Sequence 168, App
3	500.5	52.0	146	4	US-09-489-847-161 Sequence 161, App
4	500.5	52.0	146	4	US-09-904-615-121 Sequence 121, App
5	500.5	52.0	146	4	US-09-904-615-169 Sequence 169, App
6	394	41.0	79	4	US-09-904-615-167 Sequence 167, App
7	394	41.0	80	4	US-09-904-615-120 Sequence 120, App
8	85	8.8	420	4	US-09-489-039A-8961 Sequence 8961, Ap
9	80	8.3	896	4	US-09-489-039A-10262 Sequence 10262, A
10	79.5	8.3	347	4	US-09-328-352-7107 Sequence 7107, Ap
11	79.5	8.3	928	4	US-09-134-000C-5307 Sequence 5307, Ap
12	76	7.9	271	3	US-09-077-675A-12 Sequence 12, Appl
13	76	7.9	271	4	US-09-077-674-12 Sequence 12, Appl
14	76	7.9	302	3	US-09-077-675A-7 Sequence 7, Appli
15	76	7.9	302	4	US-09-077-674-7 Sequence 8, Appli
16	76	7.9	361	3	US-09-077-675A-8 Sequence 8, Appli
17	76	7.9	361	4	US-09-077-674-8 Sequence 8, Appli
18	76	7.9	366	3	US-09-077-675A-13 Sequence 13, Appl
19	76	7.9	366	4	US-09-077-674-13 Sequence 13, Appl
20	76	7.9	366	4	US-09-170-496D-88 Sequence 88, Appl
21	76	7.9	366	4	US-09-743-742B-7 Sequence 7, Appli
22	76	7.9	366	4	US-09-762-661A-5 Sequence 5, Appli
23	76	7.9	366	4	US-09-364-425B-45 Sequence 45, Appl
24	76	7.9	366	4	US-09-743-475-4 Sequence 4, Appli
25	75.5	7.8	459	4	US-09-252-991A-32058 Sequence 32058, A
26	75.5	7.8	1275	3	US-09-120-513-2 Sequence 2, Appli
27	75.5	7.8	1275	3	US-09-450-105-2 Sequence 2, Appli

28	75	7.8	302	3	US-09-077-675A-2	Sequence 2, Appli
29	75	7.8	302	4	US-09-077-674-2	Sequence 2, Appli
30	75	7.8	353	3	US-09-077-675A-3	Sequence 3, Appli
31	75	7.8	353	4	US-09-077-674-3	Sequence 3, Appli
32	75	7.8	366	4	US-09-762-661A-7	Sequence 7, Appli
33	75	7.8	366	4	US-09-743-475-6	Sequence 6, Appli
34	74.5	7.7	284	4	US-09-165-827C-14	Sequence 14, Appl
35	74.5	7.7	342	4	US-09-165-827C-2	Sequence 2, Appli
36	74.5	7.7	571	3	US-08-961-083-4	Sequence 4, Appli
37	74.5	7.7	571	4	US-09-536-784-4	Sequence 4, Appli
38	74.5	7.7	1964	2	US-08-790-912-3	Sequence 3, Appli
39	74.5	7.7	2052	2	US-08-790-912-2	Sequence 2, Appli
40	73.5	7.6	1280	4	US-09-672-810-2	Sequence 2, Appli
41	73.5	7.6	1283	4	US-09-672-810-4	Sequence 4, Appli
42	73	7.6	294	4	US-09-424-349A-2	Sequence 2, Appli
43	73	7.6	294	4	US-09-424-349A-9	Sequence 9, Appli
44	72.5	7.5	855	2	US-09-027-337-2	Sequence 2, Appli
45	72.5	7.5	855	4	US-09-644-600-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-09-904-615-114
; Sequence 114, Application US/09904615
; Patent No. 6566325
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 49 Human Secreted Proteins
; FILE REFERENCE: P2032P1
; CURRENT APPLICATION NUMBER: US/09/904,615
; CURRENT FILING DATE: 2001-07-16
; PRIOR APPLICATION NUMBER: 09/511,554
; PRIOR FILING DATE: 2000-02-23
; PRIOR APPLICATION NUMBER: 60/097,917
; PRIOR FILING DATE: 1998-08-25
; PRIOR APPLICATION NUMBER: 60/098,634
; PRIOR FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 170
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 114
; LENGTH: 182
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-904-615-114

Query Match	100.0%	Score 962;	DB 4;	Length 182;
Best Local Similarity	100.0%;	Pred. No. 4.4e-105;		
Matches 182;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MEPQLGPEAAALPGWLALLLWVSALSCSFSLPASSLSSLVPOVRTSYNFGRTFLGDKC	60	
Db	1	MEPQLGPEAAALPGWLALLLWVSALSCSFSLPASSLSSLVPOVRTSYNFGRTFLGDKC	60	
QY	61	NACIGTSICKKFFKEIRSDNWLASHLGLPPDSLLSYPNYSDDSKIWRPVEIFRLVSKY	120	
Db	61	NACIGTSICKKFFKEIRSDNWLASHLGLPPDSLLSYPNYSDDSKIWRPVEIFRLVSKY	120	
QY	121	ONEISDRKICASAPKTCISIERVLRKTERFQKWLQAKRLTPDLVQDCHQGRELKFLCM	180	
Db	121	ONEISDRKICASAPKTCISIERVLRKTERFQKWLQAKRLTPDLVQDCHQGRELKFLCM	180	
QY	181	LR 182		
Db	181	LR 182		

RESULT 2
US-09-904-615-168
; Sequence 168, Application US/09904615
; Patent No. 6566325
; GENERAL INFORMATION:

```
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 49 Human Secreted Proteins
; FILE REFERENCE: PZ032P1
; CURRENT APPLICATION NUMBER: US/09/904,615
; PRIOR FILING DATE: 2001-07-16
; PRIOR APPLICATION NUMBER: 09/511,554
; PRIOR FILING DATE: 2000-02-23
; PRIOR APPLICATION NUMBER: 60/097,917
; PRIOR FILING DATE: 1998-08-25
; PRIOR APPLICATION NUMBER: 60/098,634
; PRIOR FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 170
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 168
; LENGTH: 209
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-904-615-168

Query Match      100.0%; Score 962; DB 4; Length 209;
Best Local Similarity 100.0%; Pred. No. 5.4e-105;
Matches 182; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEPQLGPEAAALRPGWLALLLVVSALSCSFSLPASSLSLVPQVRTSYNFGRTFLGLDKC 60
Db 28 MEPQLGPEAAALRPGWLALLLVVSALSCSFSLPASSLSLVPQVRTSYNFGRTFLGLDKC 87
QY 61 NACIGTSICKKFFKEEIRSDNWLASHLGLPPDLSLLSYPPANYSDDSKIWRPVEIFRLVSKY 120
Db 88 NACIGTSICKKFFKEEIRSDNWLASHLGLPPDLSLLSYPPANYSDDSKIWRPVEIFRLVSKY 147
QY 121 QNEISDRKICASASAPKTCISIERVLRKTERFQKWLQAKRLTPDLVQDCHQGQRELKFLCM 180
Db 148 QNEISDRKICASASAPKTCISIERVLRKTERFQKWLQAKRLTPDLVQDCHQGQRELKFLCM 207
QY 181 LR 182
Db 208 LR 209

RESULT 3
US-09-489-847-161
; Sequence 161, Application US/09489847
; Patent No. 6476195
; GENERAL INFORMATION:
; APPLICANT: Rosen et al
; TITLE OF INVENTION: 98 Human Secreted Proteins
; FILE REFERENCE: PZ031P1
; CURRENT APPLICATION NUMBER: US/09/489,847
; CURRENT FILING DATE: 2000-01-24
; EARLIER APPLICATION NUMBER: PCT/US99/17130
; EARLIER FILING DATE: 1999-07-29
; EARLIER APPLICATION NUMBER: 60/094,657
; EARLIER FILING DATE: 1998-07-30
; EARLIER APPLICATION NUMBER: 60/095,486
; EARLIER FILING DATE: 1998-08-05
; EARLIER APPLICATION NUMBER: 60/096,319
; EARLIER FILING DATE: 1998-08-12
; EARLIER APPLICATION NUMBER: 60/095,454
; EARLIER FILING DATE: 1998-08-06
; EARLIER APPLICATION NUMBER: 60/095,455
; EARLIER FILING DATE: 1998-08-06
; NUMBER OF SEQ ID NOS: 376
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 161
; LENGTH: 146
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (96)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; FEATURE:
```

```
; NAME/KEY: SITE
; LOCATION: (107)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (111)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (115)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (122)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (132)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-489-847-161

Query Match      52.0%; Score 500.5; DB 4; Length 146;
Best Local Similarity 72.4%; Pred. No. 6.4e-51;
Matches 105; Conservative 4; Mismatches 33; Indels 3; Gaps 3;

QY 1 MEPQLGPEAAALRPGWLALLLVVSALSCSFSLPASSLSLVPQVRTSYNFGRTFLGLDKC 60
Db 1 MEPQLGPEAAALRPGWLALLLVVSALSCSFSLPASSLSLVPQVRTSYNFGRTFLGLDKC 60
QY 61 NACIGTSICKKFFKEEIRSDNWLASHLGLPPDLSLLSYPPANYSDDLKIWRPVEIFRLVLS 118
Db 61 NACIGTSICKKFFKEEIRSDNWLASHLGLTASRFLXSYPCKLLQIMIXKIWPXCGXLLTGQ 120
QY 119 KYQNEISDRKI-CASASAPKTCISIE 142
Db 121 QXSNEISKQELXCLJHPPPKNLHID 145

RESULT 4
US-09-904-615-121
; Sequence 121, Application US/09904615
; Patent No. 6566325
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 49 Human Secreted Proteins
; FILE REFERENCE: PZ032P1
; CURRENT APPLICATION NUMBER: US/09/904,615
; CURRENT FILING DATE: 2001-07-16
; PRIOR APPLICATION NUMBER: 09/511,554
; PRIOR FILING DATE: 2000-02-23
; PRIOR APPLICATION NUMBER: 60/097,917
; PRIOR FILING DATE: 1998-08-25
; PRIOR APPLICATION NUMBER: 60/098,634
; PRIOR FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 170
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 121
; LENGTH: 146
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (96)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (107)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (111)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (115)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
```


NAME/KEY: SITE
LOCATION: (122)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (132)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-904-615-121

Query Match 52.0%; Score 500.5; DB 4; Length 146;
Best Local Similarity 72.4%; Pred. No. 6.4e-51;
Matches 105; Conservative 4; Mismatches 33; Indels 3; Gaps 3;

QY 1 MEPQLGPEAAALRPGWLALLWVSALSCSFLPASSLSLVPQVTSYNFGRFTFLGLDKC 60
Db 1 MEPQLGPEAAALRPGWLALLWVSALSCSFLPASSLSLVPQVTSYNFGRFTFLGLDKC 60
QY 61 NACIGTSICKKFFKEEIRSDNWLASHLGLPPD-SLLSYPNYSD-DSKIWRPVEIFRLVS 118
Db 61 NACIGTSICKKFFKEEIRSDNWLASHLGLTASRFPLXSYPCKLLQMIKXKWXPCGXLLTGQ 120
QY 119 KYQNEISDRKI-CASASAPKTCISIE 142
Db 121 QXSNEISKQEIIXCLLHPPPKNLHID 145

RESULT 5

US-09-904-615-169
Sequence 169, Application US/09904615
Patent No. 6566325
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: 49 Human Secreted Proteins
FILE REFERENCE: P2032P1
CURRENT APPLICATION NUMBER: US/09/904,615
CURRENT FILING DATE: 2001-07-16
PRIOR APPLICATION NUMBER: 09/511,554
PRIOR FILING DATE: 2000-02-23
PRIOR APPLICATION NUMBER: 60/097,917
PRIOR FILING DATE: 1998-08-25
PRIOR APPLICATION NUMBER: 60/098,634
PRIOR FILING DATE: 1998-08-31
NUMBER OF SEQ ID NOS: 170
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 169
LENGTH: 146
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (96)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (107)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (111)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (115)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (122)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (132)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-904-615-169

Query Match 52.0%; Score 500.5; DB 4; Length 146;
Best Local Similarity 72.4%; Pred. No. 6.4e-51;
Matches 105; Conservative 4; Mismatches 33; Indels 3; Gaps 3;

QY 1 MEPQLGPEAAALRPGWLALLWVSALSCSFLPASSLSLVPQVTSYNFGRFTFLGLDKC 60

Db 1 MEPQLGPEAAALRPGWLALLWVSALSCSFLPASSLSLVPQVTSYNFGRFTFLGLDKC 60
QY 61 NACIGTSICKKFFKEEIRSDNWLASHLGLPPD-SLLSYPNYSD-DSKIWRPVEIFRLVS 118
Db 61 NACIGTSICKKFFKEEIRSDNWLASHLGLTASRFPLXSYPCKLLQMIKXKWXPCGXLLTGQ 120
QY 119 KYQNEISDRKI-CASASAPKTCISIE 142
Db 121 QXSNEISKQEIIXCLLHPPPKNLHID 145

RESULT 6

US-09-904-615-167
Sequence 167, Application US/09904615
Patent No. 6566325
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: 49 Human Secreted Proteins
FILE REFERENCE: P2032P1
CURRENT APPLICATION NUMBER: US/09/904,615
CURRENT FILING DATE: 2001-07-16
PRIOR APPLICATION NUMBER: 09/511,554
PRIOR FILING DATE: 2000-02-23
PRIOR APPLICATION NUMBER: 60/097,917
PRIOR FILING DATE: 1998-08-25
PRIOR APPLICATION NUMBER: 60/098,634
PRIOR FILING DATE: 1998-08-31
NUMBER OF SEQ ID NOS: 170
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 167
LENGTH: 79
TYPE: PRT
ORGANISM: Homo sapiens
US-09-904-615-167

Query Match 41.0%; Score 394; DB 4; Length 79;
Best Local Similarity 100.0%; Pred. No. 9.1e-39;
Matches 75; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEPQLGPEAAALRPGWLALLWVSALSCSFLPASSLSLVPQVTSYNFGRFTFLGLDKC 60
Db 1 MEPQLGPEAAALRPGWLALLWVSALSCSFLPASSLSLVPQVTSYNFGRFTFLGLDKC 60
QY 61 NACIGTSICKKFFKE 75
Db 61 NACIGTSICKKFFKE 75

RESULT 7

US-09-904-615-120
Sequence 120, Application US/09904615
Patent No. 6566325
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: 49 Human Secreted Proteins
FILE REFERENCE: P2032P1
CURRENT APPLICATION NUMBER: US/09/904,615
CURRENT FILING DATE: 2001-07-16
PRIOR APPLICATION NUMBER: 09/511,554
PRIOR FILING DATE: 2000-02-23
PRIOR APPLICATION NUMBER: 60/097,917
PRIOR FILING DATE: 1998-08-25
PRIOR APPLICATION NUMBER: 60/098,634
PRIOR FILING DATE: 1998-08-31
NUMBER OF SEQ ID NOS: 170
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 120
LENGTH: 80
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE

```
; LOCATION: (80)
; OTHER INFORMATION: Xaa equals stop translation
US-09-904-615-120

Query Match      41.0%; Score 394; DB 4; Length 80;
Best Local Similarity 100.0%; Pred. No. 9.2e-39;
Matches 75; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEPQLGPEAAALRPGWLALLLWVSALSCSFSLPASSLSSLPQVRTSYNFGRTFLGLDKC 60
   |||||
Db 1 MEPQLGPEAAALRPGWLALLLWVSALSCSFSLPASSLSSLPQVRTSYNFGRTFLGLDKC 60
   |||||
QY 61 NACIGTSICKKFFKE 75
   |||||
Db 61 NACIGTSICKKFFKE 75
   |||||

RESULT 8
US-09-489-039A-8961
; Sequence 8961, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/39/489,039A
; PRIOR FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 8961
; LENGTH: 420
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-8961

Query Match      8.8%; Score 85; DB 4; Length 420;
Best Local Similarity 22.2%; Pred. No. 0.22;
Matches 51; Conservative 31; Mismatches 60; Indels 88; Gaps 13;

QY 7 PEAAALRPGWLALLLWVSALSCSFSLPASSLSSLPQVRTSYNFGRTFLG 56
   |||||
Db 81 PQEIAAAGAVVWSL-----CSTDETIEEAKEKLP-----RNLCPLIKSSYGFQGT--- 127
   |||||

QY 57 LDKC-----NACIGTSIC---KKFFKEEIRSDNWLASHLGLPPDLSYPANYSDDSKI 107
   |||||
Db 128 -DKCPYFYFSDLVVGTTCDGKKWYE---YMAEFKAVHVMQLPNS-----ASDAASRAL 178
   |||||

QY 138 KRPVEIERL-----VSKYQNEISDRKI----- 129
   |||||
Db 179 WK-TEILRLQOVIEARFGTFISEAALREAIVLKNRERRALAHFYRLGQLNPPALSGDIL 237
   |||||

QY 130 -----CASASAPKTCISIERVLKTERF-QKWLOAKRLTPD---LVQDCHQG 171
   |||||
Db 238 KVVYGATFRFDKLTALIDELHAMAERIHQEWQQGKRLEPRRILITGCPIG 287
   |||||

RESULT 9
US-09-489-039A-10262
; Sequence 10262, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 10262
```

```
; LENGTH: 896
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-10262

Query Match      8.3%; Score 80; DB 4; Length 896;
Best Local Similarity 23.4%; Pred. No. 2.6;
Matches 34; Conservative 18; Mismatches 57; Indels 36; Gaps 6;

QY 6 GPEAAALRPGWLALLLWVSALSCSFSLPASSLSSLPQVRTSYNFG---RTFLGL 57
   |||||
Db 269 GQALAAARGW-----DVAMVAAPCGLPADAVAGFYREFIAAPRAITLYTMGINQSASGS 323
   |||||

QY 58 DKNACIGTSI-----CKKFF-----KKEIRSDNWLASHLGLPPDLSLSYP 98
   |||||
Db 324 DKNALINVHLASGKYGRRGCGPFSLTGQPNAMGGREVGGLATMLAAHMDVFPDDLQRLA 383
   |||||

QY 99 ANYSDDSKIWRP-----VEIFRLVSK 119
   |||||
Db 384 RFWGTERLAQTPGLTAVELFAAIGR 408
   |||||

RESULT 10
US-09-328-352-7107
; Sequence 7107, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTE
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 7107
; LENGTH: 347
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-7107

Query Match      8.3%; Score 79.5; DB 4; Length 347;
Best Local Similarity 25.5%; Pred. No. 0.76;
Matches 25; Conservative 19; Mismatches 37; Indels 17; Gaps 4;

QY 5 LGP-EAALRPGWLALLLWVSALSCSFSLP-----ASSLSSLPQVRTSYNFGRTF 54
   |||||
Db 181 LAPLSGVALSPNELVKILFNGITLQGAVPITNLGNPKKAASVRANTPAIEAQYQFGKS- 239
   |||||

QY 55 LGLDKCNACIGTSICKKFF-----KKEIRSDNWLASHL 87
   |||||
Db 240 -GVNKRFPYLGVLGYAHFNDIKLNDEIRSDLSAGHM 276
   |||||

RESULT 11
US-09-134-000C-5307
; Sequence 5307, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lytyn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5307
; LENGTH: 928
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-134-000C-5307
```


Db 157 SIRDQNEKQTVKMLAVVVFAPILCWLPFHVGRYLFKSPFPGSLEIAQISQYCNLVSFVL 216
QY 96 SY-----PANYSDDSKIWRPVEIFRLVSKYQNEISDRKI 129
Db 217 FYLSAAINPILYNIMSKKYR-VAVFRLLG--FEPFQSQRKL 253

RESULT 14
US-09-077-675A-7
; Sequence 7, Application US/09077675A
; Patent No. 6242199
; GENERAL INFORMATION:
; APPLICANT: Pai, Lee-Yuh
; APPLICANT: Feighner, Scott C.
; APPLICANT: Howard, Andrew D.
; APPLICANT: Pong, Sheng-Shung
; APPLICANT: Van Der Ploeg, Leonardus H.T.
; TITLE OF INVENTION: RECEPTOR ASSAY
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merck & Co., Inc.
; STREET: P.O. Box 2000, 126 E. Lincoln Ave.
; CITY: Rahway
; STATE: NJ
; COUNTRY: USA
; ZIP: 07065-0900
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/077,675A
; FILING DATE: 3-JUN-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Cocuzzo, Anna L.
; REGISTRATION NUMBER: 42,452
; REFERENCE/DOCKET NUMBER: 19590P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 732-594-1273
; TELEFAX: 732-594-4720
; TELEX:
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 302 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-077-675A-7

Query Match 7.9%; Score 76; DB 3; Length 302;
Best Local Similarity 25.6%; Pred. No. 1.6;
Matches 41; Conservative 20; Mismatches 51; Indels 48; Gaps 8;
QY 7 PEAAALRPGNLALLMVSALSCFSFLPASSLSLVLPQVRTSYNFGRTFLGLDKCNACIGT 66
Db 136 PTEFAVRSGLLTVMWVS--SIFFFLPVFCCLTVLYSLI-----GRKLWRRRGDAVGA 187
QY 67 SICKKFFKEEIRSDN-----WLASHLG-----LPPDSL-----L 95
Db 188 SLRDQNHKQTVKMLAVVVFAPILCWLPFHVGRYLFKSPFPGSLEIAQISQYCNLVSFVL 247
QY 96 SY-----PANYSDDSKIWRPVEIFRLVSKYQNEISDRKI 129
Db 248 FYLSAAINPILYNIMSKKYR-VAVFRLLG--FEPFQSQRKL 284

RESULT 15
US-09-077-674-7
; Sequence 7, Application US/09077674
; Patent No. 6531314
; GENERAL INFORMATION:
; APPLICANT: Arena, Joseph P.
; APPLICANT: Cully, Doris F.
; APPLICANT: Feighner, Scott D.
; APPLICANT: Howard, Andrew D.
; APPLICANT: Liberator, Paul A.
; APPLICANT: Schaeffer, James M.
; APPLICANT: Van Der Ploeg, Leonardus
; TITLE OF INVENTION: GROWTH HORMONE SECRETAGOGUE RECEPTOR FAMILY
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merck & Co., Inc.
; STREET: P.O. Box 2000, 126 E. Lincoln Ave.
; CITY: Rahway
; STATE: NJ
; COUNTRY: USA
; ZIP: 07065-0900
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/077,674
; FILING DATE: 3-JUN-1998
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Cocuzzo, Anna L.
; REGISTRATION NUMBER: 42,452
; REFERENCE/DOCKET NUMBER: 19589P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 732-594-1273
; TELEFAX: 732-594-4720
; TELEX:
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 302 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-077-674-7

Query Match 7.9%; Score 76; DB 4; Length 302;
Best Local Similarity 25.6%; Pred. No. 1.6;
Matches 41; Conservative 20; Mismatches 51; Indels 48; Gaps 8;
QY 7 PEAAALRPGNLALLMVSALSCFSFLPASSLSLVLPQVRTSYNFGRTFLGLDKCNACIGT 66
Db 136 PTEFAVRSGLLTVMWVS--SIFFFLPVFCCLTVLYSLI-----GRKLWRRRGDAVGA 187
QY 67 SICKKFFKEEIRSDN-----WLASHLG-----LPPDSL-----L 95
Db 188 SLRDQNHKQTVKMLAVVVFAPILCWLPFHVGRYLFKSPFPGSLEIAQISQYCNLVSFVL 247
QY 96 SY-----PANYSDDSKIWRPVEIFRLVSKYQNEISDRKI 129
Db 248 FYLSAAINPILYNIMSKKYR-VAVFRLLG--FEPFQSQRKL 284

Search completed: June 8, 2004, 16:26:55
Job time : 23 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 8, 2004, 16:20:14 ; Search time 57 Seconds
(without alignments)
902.169 Million cell updates/sec

Title: US-10-054-988-114
Perfect score: 962
Sequence: 1 MEPQLGPERAALRPGWLALL.....DLVQDCEQSQRELKFLCMLR 182

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: genesecp1980s:*
2: genesecp1990s:*
3: genesecp2000s:*
4: genesecp2001s:*
5: genesecp2002s:*
6: genesecp2003as:*
7: genesecp2003bs:*
8: genesecp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	962	100.0	182	3	AAY91393 Human sec
2	962	100.0	182	4	AAB95695 Human pro
3	962	100.0	182	6	ABU03563 Angiogene
4	962	100.0	182	6	ABR47459 Breast ca
5	962	100.0	209	3	AAY91447 Human sec
6	959	99.7	182	4	AAU12257 Human PRO
7	959	99.7	182	4	AAB48066 Human ext
8	959	99.7	182	6	ABO17701 Novel hum
9	959	99.7	182	6	ABU80955 Human PRO
10	959	99.7	182	6	ABU66655 Human PRO
11	959	99.7	182	6	ABU59736 Novel sec
12	959	99.7	182	6	ABO24926 Human sec
13	959	99.7	182	6	ABU66931 Human sec
14	959	99.7	182	6	ADA45691 Novel hum
15	959	99.7	182	6	ADA76122 Human PRO
16	959	99.7	182	6	ADA18772 Human PRO
17	959	99.7	182	6	ADA61395 Homo sapi
18	959	99.7	182	6	ADB19180 Novel hum
19	959	99.7	182	6	ADB27721 Human PRO
20	959	99.7	182	6	ADA86200 Novel hum
21	959	99.7	182	6	ADB15764 Human PRO
22	959	99.7	182	6	ADA47550 Human PRO
23	959	99.7	182	6	ADA67345 Human PRO
24	959	99.7	182	6	ADB30352 Human PRO
25	959	99.7	182	6	ADA85648 Novel hum

26	959	99.7	182	6	ADA96860	Human PRO
27	959	99.7	182	6	ADA79164	Human PRO
28	959	99.7	182	6	ADA87303	Novel hum
29	959	99.7	182	6	ADB16505	Human PRO
30	959	99.7	182	6	ADA91597	Novel hum
31	959	99.7	182	6	ADB14660	Human PRO
32	959	99.7	182	6	ADB18621	Novel hum
33	959	99.7	182	6	ADA93836	Human PRO
34	959	99.7	182	6	ADB19732	Novel hum
35	959	99.7	182	6	ADB13044	Human PRO
36	959	99.7	182	6	ABO43234	Novel hum
37	959	99.7	182	6	ADA74298	Human PRO
38	959	99.7	182	6	ADB24531	Human PRO
39	959	99.7	182	6	ADA82055	Human PRO
40	959	99.7	182	6	ADA75018	Human PRO
41	959	99.7	182	6	ADA85096	Novel hum
42	959	99.7	182	6	ADA84544	Novel hum
43	959	99.7	182	6	ADB29800	Human PRO
44	959	99.7	182	6	ADA80328	Human PRO
45	959	99.7	182	6	ADA75570	Human PRO

ALIGNMENTS

RESULT 1
AAY91393
ID AAY91393 standard; protein; 182 AA.
XX AAY91393;
AC AAY91393;
XX
DT 29-JUN-2000 (first entry)
XX
DE Human secreted protein sequence encoded by gene 48 SEQ ID NO:114.
XX
KW Human; secreted protein; diagnosis; neuroprotective; nootropic;
KW neuroleptic; antimanic; cerebroprotective; immunomodulatory;
KW anti-microbial; cardiant; cytostatic; antiinflammatory; haemostatic;
KW anticonvulsant; vasotropic; vaccine; gene therapy; anti-sense therapy;
KW neural; reproductive; immune disorder; immunodeficiency; infection;
KW lymphoma; demyelinating disease; autoimmunity; cancer; inflammation;
KW aneurysm; haemorrhage; Alzheimer's disease; Parkinson's disease;
KW Huntington's disease; Tourette syndrome; multiple sclerosis; meningitis;
KW ischaemia; mania; dementia; obsessive compulsive disorder;
KW viral prophylaxis; developmental disorder; sexually-linked disorder;
KW cardiovascular disorder; food additive; preservative.
XX
OS Homo sapiens.
XX
PN WO200011014-A1.
XX
PD 02-MAR-2000.
XX
PF 24-AUG-1999; 99WO-US019330.
XX
PR 25-AUG-1998; 98US-0097917P.
PR 31-AUG-1998; 98US-0098634P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Moore PA, Ruben SM, Olsen HS, Shi Y, Rosen CA, Florence KA;
PI Soppet DR, Lafleur DW, Endress GA, Ebner R, Komatsoulis G, Duan RD;
XX
DR WPI; 2000-224656/19.
DR N-PSDB; AAA26328.
XX
PT Novel secreted proteins and corresponding DNA molecules that can be used
PT to prevent, treat and diagnose disease in humans, for example,
PT Alzheimer's, cancer, and immune disorders.
XX
PS Claim 11; Page 380-381; 416pp; English.
XX
CC The polynucleotide sequences given in AAA26281 to AAA26336 encode the

CC human secreted proteins given in AAY91346 to AAY91449. The human secreted
CC proteins can have activities based on the tissues and cells they are
CC expressed in. Examples of the activities are: neuroprotective; nootropic;
CC neuroleptic; antitumor; cerebroprotective; immunomodulatory; anti-
CC microbial; cardiant; cytostatic; antiinflammatory; haemostatic;
CC anticonvulsant; and vasotrophic. The polynucleotides and proteins may be
CC used to prevent, treat or ameliorate a medical condition, e.g. by protein
CC or gene therapy. Conditions treatable by the proteins of the invention
CC include neural, reproductive, or immune disorders, especially
CC immunodeficiency, infection, lymphomas, demyelinating diseases, auto-
CC immunities, cancer, general microbial infection, inflammation, aneurysms
CC and haemorrhages. Specific examples include: Alzheimer's disease;
CC Parkinson's; Huntington's; Tourette syndrome; mania; dementia; obsessive
CC compulsive disorder and viral prophylaxis. The polynucleotides and
CC proteins can also be used in the detection of disorders associated with
CC the function of the protein, for example, the detection of developmental
CC disorders, sexually-linked disorders, or disorders of the cardiovascular
CC system. They may also be used as food additives or preservatives.
CC AAA26272 to AAA26280 and AAY91345 are sequences used in the
CC exemplification of the present invention
XX

Sequence 182 AA;

Query Match 100.0%; Score 962; DB 3; Length 182;
Best Local Similarity 100.0%; Pred. No. 3.6e-95;
Matches 182; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEPQLGPEAAALRPGWLALLLWVSALSCSFSLPSSLSLVPQVRTSYNFGRTFLGLDKC 60
Db |||||
1 MEPQLGPEAAALRPGWLALLLWVSALSCSFSLPSSLSLVPQVRTSYNFGRTFLGLDKC 60

QY 61 NACIGTSICKKFFKEIRSDNWLASHLGLPPDSLLSYPPANYSDDSKIWRPVEIFRLVSKY 120
Db |||||
61 NACIGTSICKKFFKEIRSDNWLASHLGLPPDSLLSYPPANYSDDSKIWRPVEIFRLVSKY 120

QY 121 QNEISDRKICASAPKTCISIERVLRKTERFQKWLQAKRLTPDLVQDCHQORELKFCLM 180
Db |||||
121 QNEISDRKICASAPKTCISIERVLRKTERFQKWLQAKRLTPDLVQDCHQORELKFCLM 180

QY 181 LR 182
Db ||
181 LR 182

RESULT 2
AAB95695
ID AAB95695 standard; protein; 182 AA.
XX

AC AAB95695;

XX 26-JUN-2001 (first entry)

DE Human protein sequence SEQ ID NO:18516.

XX Human; primer; detection; diagnosis; antisense therapy; gene therapy.

OS Homo sapiens.

XX EP1074617-A2.

PN 07-FEB-2001.

PD 28-JUL-2000; 2000EP-00116126.

XX 29-JUL-1999; 99JP-00248036.

PR 27-AUG-1999; 99JP-00300253.

PR 11-JAN-2000; 2000JP-00118776.

PR 02-MAY-2000; 2000JP-00183767.

PR 09-JUN-2000; 2000JP-00241899.

XX (HELI-) HELIX RES INST.

XX

PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX WPI; 2001-318749/34.
DR
XX
PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-
PT length cDNAs defined in the specification, and for the detection and/or
PT diagnosis of the abnormality of the proteins encoded by the full-length
PT cDNAs.

PS Claim 8; SEQ ID NO 18516; 2537pp + Sequence Listing; English.

XX The present invention describes primer sets for synthesizing 5602 full-
CC length cDNAs defined in the specification. Where a primer set comprises:
CC (a) an oligo-dT primer and an oligonucleotide complementary to the
CC complementary strand of a polynucleotide which comprises one of the 5602
CC nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in the
CC specification. The primer sets can be used in antisense therapy and in
CC gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
CC oligonucleotides, all of which are used in the exemplification of the
CC present invention
XX

Sequence 182 AA;

Query Match 100.0%; Score 962; DB 4; Length 182;
Best Local Similarity 100.0%; Pred. No. 3.6e-95;
Matches 182; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEPQLGPEAAALRPGWLALLLWVSALSCSFSLPSSLSLVPQVRTSYNFGRTFLGLDKC 60
Db |||||
1 MEPQLGPEAAALRPGWLALLLWVSALSCSFSLPSSLSLVPQVRTSYNFGRTFLGLDKC 60

QY 61 NACIGTSICKKFFKEIRSDNWLASHLGLPPDSLLSYPPANYSDDSKIWRPVEIFRLVSKY 120
Db |||||
61 NACIGTSICKKFFKEIRSDNWLASHLGLPPDSLLSYPPANYSDDSKIWRPVEIFRLVSKY 120

QY 121 QNEISDRKICASAPKTCISIERVLRKTERFQKWLQAKRLTPDLVQDCHQORELKFCLM 180
Db |||||
121 QNEISDRKICASAPKTCISIERVLRKTERFQKWLQAKRLTPDLVQDCHQORELKFCLM 180

QY 181 LR 182
Db ||
181 LR 182

RESULT 3
ABU03563
ID ABU03563 standard; protein; 182 AA.
XX

AC ABU03563;

XX 21-JAN-2003 (first entry)

DE Angiogenesis-associated human protein sequence #108.

XX Human; angiogenesis-associated transcript; angiogenesis;

KW angiogenesis-associated disease; cancer; cytostatic.

XX Homo sapiens.

OS

PN WO200279492-A2.
XX 10-OCT-2002.
XX 14-FEB-2002; 2002WO-US004915.
XX 14-FEB-2001; 2001US-00784356.
PR 22-FEB-2001; 2001US-00791390.
PR 19-APR-2001; 2001US-0285475P.
PR 03-AUG-2001; 2001US-0310025P.
PR 13-NOV-2001; 2001US-0350666P.
PR 29-NOV-2001; 2001US-0334244P.
XX (EOSB-) EOS BIOTECHNOLOGY INC.
XX Murray R, Glynn R, Watson SR, Aziz N;
XX WPI; 2003-040681/03.
DR N-PSDB; ABX08847.
XX
PT Detecting angiogenesis-associated transcript in a cell for diagnosing and
PT treating cancer by contacting a sample with a polynucleotide that
PT exhibits changes in expression level as a function of time in tissue
PT undergoing angiogenesis.
XX
PS Example 2; Page 282; 291pp; English.
XX
CC The present invention relates to methods and compositions for detecting
CC an angiogenesis-associated transcript in a cell in a patient. The method
CC involves contacting a biological sample from the patient with a
CC polynucleotide that selectively hybridizes to a sequence at least 80%
CC identical to any of the angiogenesis-associated human polynucleotide
CC sequences given in the specification. These angiogenesis-associated
CC polynucleotide sequences comprise genes that exhibit changes in
CC expression levels as a function of time in tissue undergoing
CC angiogenesis. The method and the polynucleotide sequences of the
CC invention are useful for diagnosing and treating angiogenesis and
CC angiogenesis-associated diseases e.g. cancer. The polynucleotide
CC sequences are also useful in the gene therapy of such disorders. The
CC angiogenesis-associated proteins encoded by the polynucleotide sequences
CC are useful as a vaccine for therapeutic and prophylactic immunisation.
CC ABU03456-ABU03569 represent angiogenesis-associated protein sequences
XX
SQ Sequence 182 AA;
Query Match 100.0%; Score 962; DB 6; Length 182;
Best Local Similarity 100.0%; Pred. No. 3.6e-95;
Matches 182; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MEPQLGPEAAALRPGWLALLWVSALSCSFSLPASSLSLVPQVTSYNFGRFTGLDKC 60
Db 1 MEPQLGPEAAALRPGWLALLWVSALSCSFSLPASSLSLVPQVTSYNFGRFTGLDKC 60
QY 61 NACIGTSICKFFKEEIRSDNWLASHLGLPPDLSLLSYYPANYSDDSKIWRPVEIFRLVSKY 120
Db 61 NACIGTSICKFFKEEIRSDNWLASHLGLPPDLSLLSYYPANYSDDSKIWRPVEIFRLVSKY 120
QY 121 QNEISDRKICASAPKTCSTIERVLKTERFQKWLQAKRLTPDLVQDCHQORELKFLCM 180
Db 121 QNEISDRKICASAPKTCSTIERVLKTERFQKWLQAKRLTPDLVQDCHQORELKFLCM 180
QY 181 LR 182
Db 181 LR 182
RESULT 4
ABR47459
ID ABR47459 standard; protein; 182 AA.
XX
AC ABR47459;
XX
DT 12-JUN-2003 (first entry)

XX Breast cancer associated protein sequence SEQ ID NO:150.
DE Human; breast cancer; cytostatic; gene therapy.
XX Homo sapiens.
XX WO2003004989-A2.
XX 16-JAN-2003.
XX 21-JUN-2002; 2002WO-US019669.
XX 21-JUN-2001; 2001US-0299887P.
PR 27-JUN-2001; 2001US-0301572P.
PR 18-JUL-2001; 2001US-0306501P.
PR 25-SEP-2001; 2001US-0325002P.
PR 05-MAR-2002; 2002US-0362585P.
PR 14-MAY-2002; 2002US-0380391P.
XX (MILL-) MILLENIUM PHARM INC.
PA Lillie J, Gannavarapu M, Glatt K, Hoersh S, Kamatkar S;
XX Mertens M, Monahan JE, Myer V, Wang Y, Xu Y, Zhao X, Meyers RE;
PI Bast RC, Hortobagyi GN, Puzsai L, Meric F, Sahin A, Mills GB;
XX WPI; 2003-210381/20.
DR N-PSDB; ACC50151.
XX
PT Breast cancer diagnosis or treatment by comparing the level of expression
PT of a marker in a patient sample with that in the control non-breast
PT cancer sample.
XX
PS Claim 1; SEQ ID NO 150; 128pp; English.
XX
CC The present invention describes a method for assessing whether a patient
CC is afflicted with breast cancer. The method comprises comparing the level
CC of expression of a marker (gene/polypeptide see ACC50076 to ACC50334 and
CC ABR47386 to ABR47632) in a patient sample and the normal level of
CC expression of the marker in a control non-breast cancer sample, where a
CC significant increase in the level of expression of the marker in the
CC patient sample and the normal level is an indication that the patient is
CC afflicted with breast cancer. The breast cancer associated sequences from
CC the present invention have cytostatic activities and can be used in gene
CC therapy. The method is useful for diagnosing and treating breast cancer.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 182 AA;
Query Match 100.0%; Score 962; DB 6; Length 182;
Best Local Similarity 100.0%; Pred. No. 3.6e-95;
Matches 182; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MEPQLGPEAAALRPGWLALLWVSALSCSFSLPASSLSLVPQVTSYNFGRFTGLDKC 60
Db 1 MEPQLGPEAAALRPGWLALLWVSALSCSFSLPASSLSLVPQVTSYNFGRFTGLDKC 60
QY 61 NACIGTSICKFFKEEIRSDNWLASHLGLPPDLSLLSYYPANYSDDSKIWRPVEIFRLVSKY 120
Db 61 NACIGTSICKFFKEEIRSDNWLASHLGLPPDLSLLSYYPANYSDDSKIWRPVEIFRLVSKY 120
QY 121 QNEISDRKICASAPKTCSTIERVLKTERFQKWLQAKRLTPDLVQDCHQORELKFLCM 180
Db 121 QNEISDRKICASAPKTCSTIERVLKTERFQKWLQAKRLTPDLVQDCHQORELKFLCM 180
QY 181 LR 182
Db 181 LR 182
RESULT 5

AAAY91447
ID AAY91447 standard; protein; 209 AA.
XX AC AAY91447;
XX DT 29-JUN-2000 (first entry)
XX DE Human secreted protein sequence encoded by gene 48 SEQ ID NO:168.
XX KW Human; secreted protein; diagnosis; neuroprotective; neurotropic;
KW neuroleptic; antimanic; cerebroprotective; immunomodulatory;
KW anti-microbial; cardiant; cytostatic; antiinflammatory; haemostatic;
KW anticonvulsant; vasotrophic; vaccine; gene therapy; anti-sense therapy;
KW neural; reproductiv; immune disorder; immunodeficiency; infection;
KW lymphoma; demyelinating disease; autoimmunity; cancer; inflammation;
KW aneurysm; haemorrhage; Alzheimer's disease; Parkinson's disease;
KW Huntington's disease; Tourette syndrome; multiple sclerosis; meningitis;
KW ischaemia; mania; dementia; obsessive compulsive disorder;
KW viral prophylaxis; developmental disorder; sexually-linked disorder;
KW cardiovascular disorder; food additive; preservative.
XX OS Homo sapiens.
XX PN WO200011014-A1.
XX PD 02-MAR-2000.
XX PF 24-AUG-1999; 99WO-US019330.
XX PR 25-AUG-1998; 98US-0097917P.
XX PR 31-AUG-1998; 98US-0098634P.
XX PA (HUMA-) HUMAN GENOME SCI INC.
XX PI Moore PA, Ruben SM, Olsen HS, Shi Y, Rosen CA, Florence KA;
PI Soppet DR, Lafleur DW, Endress GA, Ebner R, Komatsoulis G, Duan RD;
XX WPI; 2000-224656/19.
XX PT Novel secreted proteins and corresponding DNA molecules that can be used
PT to prevent, treat and diagnose disease in humans, for example,
PT Alzheimer's, cancer, and immune disorders.
XX PS Disclosure; Page 413-414; 416pp; English.
XX CC The polynucleotide sequences given in AAA26281 to AAA26336 encode the
CC human secreted proteins given in AAY91346 to AAY91449. The human secreted
CC proteins can have activities based on the tissues and cells they are
CC expressed in. Examples of the activities are: neuroprotective; neurotropic;
CC neuroleptic; antimanic; cerebroprotective; immunomodulatory; anti-
CC microbial; cardiant; cytostatic; antiinflammatory; haemostatic;
CC anticonvulsant; and vasotrophic. The polynucleotides and proteins may be
CC used to prevent, treat or ameliorate a medical condition, e.g. by protein
CC or gene therapy. Conditions treatable by the proteins of the invention
CC include neural, reproductive, or immune disorders, especially
CC immunodeficiency, infection, lymphomas, demyelinating diseases, auto-
CC immunities, cancer, general microbial infection, inflammation, aneurysms
CC and haemorrhages. Specific examples include: Alzheimer's disease;
CC Parkinson's; Huntington's; Tourette syndrome; multiple sclerosis;
CC meningitis; ischaemia; prostate cancer; mania; dementia; obsessive
CC compulsive disorder and viral prophylaxis. The polynucleotides and
CC proteins can also be used in the detection of disorders associated with
CC the function of the protein, for example, the detection of developmental
CC disorders, sexually-linked disorders, or disorders of the cardiovascular
CC system. They may also be used as food additives or preservatives.
CC AAA26272 to AAA26280 and AAY91345 are sequences used in the
XX exemplification of the present invention
SQ Sequence 209 AA;

Query Match 100.0%; Score 962; DB 3; Length 209;
Best Local Similarity 100.0%; Pred. No. 4.3e-95;
Matches 182; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEPQLGPEAAALRPGWLALLLWVSALSCSFSLPASSLSLVPQVRTSYNFGRTFLGLDKC 60
Db |||||
28 MEPQLGPEAAALRPGWLALLLWVSALSCSFSLPASSLSLVPQVRTSYNFGRTFLGLDKC 87
QY 61 NACIGTSICKKFFKBEIRSDNWLASHLGLPPDSLLSYNPANYSDDSKIMRPVEIFRLVSKY 120
Db |||||
88 NACIGTSICKKFFKBEIRSDNWLASHLGLPPDSLLSYNPANYSDDSKIMRPVEIFRLVSKY 147
QY 121 QNEISDRKKICASASAPKTCISIERVLKTERFQKWLQAKRLTPDLVQDCHQSGRELKFLCM 180
Db |||||
148 QNEISDRKKICASASAPKTCISIERVLKTERFQKWLQAKRLTPDLVQDCHQSGRELKFLCM 207
QY 181 LR 182
Db ||
208 LR 209
RESULT 6
AAU12257
ID AAU12257 standard; protein; 182 AA.
XX AC AAU12257;
XX DT 24-OCT-2001 (first entry)
XX DE Human PRO3743 polypeptide sequence.
XX KW Human secretory and transmembrane; PRO; mammalian; cancer; lung; breast;
KW prostate; cervical; tumour necrosis factor-alpha; TNF-alpha; cartilage;
KW ear; proliferation; glucose; free fatty acid; skeletal muscle; adipocyte;
KW A-peptide; factor VIIA; gene therapy.
XX OS Homo sapiens.
XX PN WO200140466-A2.
XX PD 07-JUN-2001.
XX PF 01-DEC-2000; 2000WO-US032678.
XX PR 01-DEC-1999; 99WO-US028301.
PR 01-DEC-1999; 99WO-US028634.
PR 02-DEC-1999; 99WO-US028551.
PR 02-DEC-1999; 99WO-US028564.
PR 02-DEC-1999; 99WO-US028565.
PR 09-DEC-1999; 99US-0170262P.
PR 16-DEC-1999; 99WO-US030095.
PR 20-DEC-1999; 99WO-US030911.
PR 20-DEC-1999; 99WO-US030999.
PR 30-DEC-1999; 99WO-US031243.
PR 30-DEC-1999; 99WO-US031274.
PR 05-JAN-2000; 2000WO-US000219.
PR 06-JAN-2000; 2000WO-US000277.
PR 06-JAN-2000; 2000WO-US000376.
PR 11-FEB-2000; 2000WO-US003565.
PR 18-FEB-2000; 2000WO-US004341.
PR 18-FEB-2000; 2000WO-US004342.
PR 22-FEB-2000; 2000WO-US004414.
PR 24-FEB-2000; 2000WO-US004914.
PR 24-FEB-2000; 2000WO-US005004.
PR 01-MAR-2000; 2000WO-US005601.
PR 02-MAR-2000; 2000WO-US005841.
PR 03-MAR-2000; 2000US-0187202P.
PR 10-MAR-2000; 2000WO-US006319.
PR 15-MAR-2000; 2000WO-US006884.
PR 20-MAR-2000; 2000WO-US007377.
PR 21-MAR-2000; 2000WO-US007532.
PR 30-MAR-2000; 2000WO-US008439.
PR 17-MAY-2000; 2000WO-US013705.
PR 22-MAY-2000; 2000WO-US014042.
PR 30-MAY-2000; 2000WO-US014941.
PR 02-JUN-2000; 2000WO-US015264.

PR 05-JUN-2000; 2000US-0209832P.
PR 28-JUL-2000; 2000WO-US020710.
PR 11-AUG-2000; 2000WO-US022031.
PR 23-AUG-2000; 2000WO-US023522.
PR 24-AUG-2000; 2000WO-US023328.
PR 08-NOV-2000; 2000WO-US030952.
PR 10-NOV-2000; 2000WO-US030873.
XX (GETH) GENENTECH INC.
PA Baker KP, Beresini M, Deforge L, Desnoyers L, Filvaroff E, Gao W;
PI Gerritsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S;
PI Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;
XX
DR WPI; 2001-408281/43.
DR N-PSDB; AAS21329.
XX
PT Isolated , secretory and transmembrane PRO polypeptide used to detect
PT other PRO polypeptides, link bioactive molecules to cells expressing PRO
PT polypeptides, and detect the presence of mammalian tumors e.g. lung,
PT breast, prostate, cervical.
XX
PS Claim 12; Fig 172; 813pp; English.
XX
CC AAU12172-AAU12446 represent novel human secretory and transmembrane PRO
CC polypeptides. The PRO polypeptides are useful to detect other PRO
CC polypeptides, to link bioactive molecules to cells expressing PRO
CC polypeptides, to modulate biological activities of cells expressing PRO
CC polypeptides, and to detect the presence of mammalian lung, colon,
CC breast, prostate, rectal, cervical or liver tumours by comparing PRO
CC polypeptide expression in a cell sample to that in a control sample. Some
CC of the 275 sequences are also useful to stimulate the release of tumour
CC necrosis factor-alpha (TNF-alpha) from human blood, the proliferation or
CC differentiation of chondrocytes, the proliferation or gene expression in
CC pericyte cells, the release of proteoglycans from cartilage, the
CC proliferation of inner ear utricular supporting cells or of T-
CC lymphocytes, the release of a cytokine from peripheral blood monocytes
CC (PBMCs), or the proliferation of endothelial cells. Some of the PRO
CC polypeptides may modulate glucose or free fatty acid uptake by skeletal
CC muscle cells or by adipocytes; or inhibit binding of A-peptide to factor
CC VIIA. The PRO polypeptides can be used in assays to identify molecules
CC involved in binding interactions. The polynucleotides encoding PRO
CC polypeptides can be used to generate probes, antisense RNA/DNA,
CC transgenic or knock out animals and can be used in gene therapy
XX
SQ Sequence 182 AA;
Query Match 99.7%; Score 959; DB 4; Length 182;
Best Local Similarity 99.5%; Pred. No. 7.6e-95;
Matches 181; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 MEPQLGPEAAALRPGWLALLLWVSALSCFSLPASSLSLVPQVRTSYNFGRTFLGLDKC 60
Db |||||
1 MEPQLGPEAAALRPGWLALLLWVSALSCFSLPASSLSLVPQVRTSYNFGRTFLGLDKC 60
QY 61 NACIGTSICKKFFKEEIRSDNWLASHLGLPPDSLLSYPNYSDSKIWRPVEIFRLVSKY 120
Db |||||
61 NACIGTSICKKFFKEEIRSDNWLASHLGLPPDSLLSYPNYSDSKIWRPVEIFRLVSKY 120
QY 121 QNEISDRKICASAPKTCISIERVLEKTERFQKWLQAKRLTPDLVQDCHQGQRELKFLCM 180
Db |||||:|||||
121 QNEISDRRICASAPKTCISIERVLEKTERFQKWLQAKRLTPDLVQDCHQGQRELKFLCM 180
QY 181 LR 182
Db ||
181 LR 182
RESULT 7
AAB48066
ID AAB48066 standard; protein; 182 AA.
XX
AC AAB48066;

XX 19-MAR-2001 (first entry)
DT Human extracellular signaling molecule (EXCS) (ID 5090841CD1).
XX
DE Extracellular signaling molecule; EXCS; anti-inflammatory; human;
XX immunosuppressive; cytostatic; neuroprotective; gastrointestinal;
KW viricide; antibacterial; anti-HIV; human immunodeficiency virus;
KW antiinfectility; cerebroprotective; nootropic; antiulcer; antifungal;
KW anticonvulsant; tranquilizer; neuroleptic; vasotropic; gynecological;
KW keratolytic; protozoacide; gene therapy.
XX Homo sapiens.
OS WO200070049-A2.
XX 23-NOV-2000.
PD 19-MAY-2000; 2000WO-US013975.
XX 19-MAY-1999; 99US-0134949P.
PF 15-JUL-1999; 99US-0144270P.
XX 30-JUL-1999; 99US-0146700P.
PR 04-OCT-1999; 99US-0157508P.
XX {INCY-} INCYTE GENOMICS INC.
PA Tang YT, Yue H, Lal P, Burford N, Bandman O, Baughn MR;
PI Azimzai Y, Lu DAM, Patterson C;
XX WPI; 2001-025021/03.
DR N-PSDB; AAC84302.
XX New human extracellular signaling nucleic acids and polypeptides useful
PT for diagnosing, treating and preventing infections and gastrointestinal,
PT neurological, reproductive, and autoimmune/inflammatory disorders.
XX Claim 1; Page 88-89; 114pp; English.
PS The invention provides human extracellular signaling molecules (EXCS) and
XX polynucleotides which identify and encode EXCS. EXCS can be expressed by
CC standard recombinant methodology. The amino acid and nucleic acid
CC sequences of EXCS are useful for diagnosing, treating and preventing
CC infections and gastrointestinal (peptic ulcer, dysphagia, pancreatitis),
CC neurological (e.g. epilepsy, ischemic cerebrovascular disease, stroke),
CC reproductive (infertility, ovulatory defects, endometriosis), autoimmune
CC /inflammatory (actinic keratosis, acquired immunodeficiency syndrome
CC (AIDS), Addison's disease), and cell proliferative disorders including
CC cancers (of the breast, adrenal gland, bone). They may also be used to
CC treat fatal familial insomnia, nutritional and metabolic diseases of the
CC nervous system, myopathies, mental disorders (anxiety, schizophrenia,
CC mood), as well as infections caused by parasites (malaria, leishmania,
CC trypanosoma), viral (adenovirus, coronavirus, flavivirus), bacterial
CC (e.g. pneumococcus, staphylococcus, bacillus), and fungal (aspergillus,
CC blastomycetes, dermatophytes) agents. The nucleic acids, polypeptides,
CC antagonists, agonists, pharmaceutical compositions, and antibodies may
CC also be used for treating or preventing disorders associated with
CC increased or decreased expression or activity of EXCS. EXCS
CC polynucleotides may also be used to detect and quantify gene expression
CC in biopsied tissues in which expression of EXCS may be correlated with
CC the disease, to determine presence or excess expression of EXCS, to
CC monitor regulation of EXCS levels during therapeutic intervention, to
CC detect the presence of associated disorders, as targets in microarray, to
CC generate hybridization probes, and to detect differences in gene
CC sequences among normal, carrier or affected individuals. Antibodies may
CC also be used in diagnosing disorders, in monitoring patients being
CC treated with EXCS agonists, antagonists or inhibitors. Sequences AAB48057
CC -B48082 represent the EXCS of the invention
XX
SQ Sequence 182 AA;
Query Match 99.7%; Score 959; DB 4; Length 182;
Best Local Similarity 99.5%; Pred. No. 7.6e-95;

Matches		181;	Conservative	1;	Mismatches	0;	Indels	0;	Gaps	0;
QY	1	MEPQLGPEAAALRPGWLALLLV	VSALSCSFSLPASSLSLLV	PQVRTSYN	EGRTFLGDKC	60				
Db	1	MEPQLGPEAAALRPGWLALLLV	VSALSCSFSLPASSLSLLV	PQVRTSYN	EGRTFLGDKC	60				
QY	61	NACIGTSICKFFKEEIRSDNWLASHLGLPPD	SLSYDPANYSDDSKIWRPVEIFRLVSKY	120						
Db	61	NACIGTSICKFFKEEIRSDNWLASHLGLPPD	SLSYDPANYSDDSKIWRPVEIFRLVSKY	120						
QY	121	QNEISDRKICASAPKTCISIERVLRKTERFQKWLQAKRLTPOLVQDCHQGORELKFLCM	180							
Db	121	QNEISDRRICASAPKTCISIERVLRKTERFQKWLQAKRLTPOLVQDCHQGORELKFLCM	180							
QY	181	LR	182							
Db	181	LR	182							
RESULT 8										
ABO17701										
ID	ABO17701 standard; protein; 182 AA.									
XX										
AC	ABO17701;									
XX										
DT	26-AUG-2003 (first entry)									
XX										
DE	Novel human secreted and transmembrane protein PRO3743.									
XX										
KW	Human; secreted and transmembrane protein; PRO; antiinflammatory;									
KW	antiarteriosclerotic; cardiant; anti-infertility; anti-HIV; cytostatic;									
KW	antidiabetic; gene therapy; tumour necrosis factor (TNF)-alpha release;									
KW	TNF-alpha release; cell proliferation; cell differentiation;									
KW	gene expression modulator; proteoglycan release; cytokine release;									
KW	tumour; inflammatory disease; organ failure; atherosclerosis;									
KW	cardiac injury; infertility; birth defect; premature aging; AIDS;									
KW	acquired immunodeficiency syndrome; cancer; diabetic complication;									
KW	chromosome mapping; gene mapping; pharmaceutical; diagnostic; biosensor;									
KW	bioreactor; tissue typing.									
XX										
OS	Homo sapiens.									
XX										
PN	US2003032156-A1.									
XX										
PD	13-FEB-2003.									
XX										
PF	06-MAY-2002; 2002US-00140474.									
XX										
PR	31-MAR-1997;	97WO-US005230.								
PR	12-JUN-1998;	98WO-US012456.								
PR	14-JUL-1998;	98WO-US014552.								
PR	28-AUG-1998;	98WO-US017888.								
PR	10-SEP-1998;	98WO-US018824.								
PR	14-SEP-1998;	98WO-US019093.								
PR	14-SEP-1998;	98WO-US019094.								
PR	14-SEP-1998;	98WO-US019177.								
PR	16-SEP-1998;	98WO-US019330.								
PR	17-SEP-1998;	98WO-US019437.								
PR	07-OCT-1998;	98WO-US021141.								
PR	29-OCT-1998;	98WO-US022991.								
PR	29-OCT-1998;	98WO-US022992.								
PR	20-NOV-1998;	98WO-US024855.								
PR	01-DEC-1998;	98WO-US025108.								
PR	05-JAN-1999;	99WO-US000106.								
PR	08-MAR-1999;	99WO-US005028.								
PR	10-MAR-1999;	99WO-US005190.								
PR	20-APR-1999;	99WO-US008615.								
PR	14-MAY-1999;	99WO-US010733.								
PR	02-JUN-1999;	99WO-US012252.								
PR	01-SEP-1999;	99WO-US020111.								
PR	08-SEP-1999;	99WO-US020594.								
PR	13-SEP-1999;	99WO-US020944.								
PR	15-SEP-1999;	99WO-US021090.								

PR	15-SEP-1999;	99WO-US021547.
PR	05-OCT-1999;	99WO-US023089.
PR	29-NOV-1999;	99WO-US028214.
PR	30-NOV-1999;	99WO-US028313.
PR	30-NOV-1999;	99WO-US028409.
PR	01-DEC-1999;	99WO-US028301.
PR	01-DEC-1999;	99WO-US028634.
PR	02-DEC-1999;	99WO-US028551.
PR	02-DEC-1999;	99WO-US028564.
PR	02-DEC-1999;	99WO-US028565.
PR	16-DEC-1999;	99WO-US030095.
PR	20-DEC-1999;	99WO-US030911.
PR	20-DEC-1999;	99WO-US030999.
PR	22-DEC-1999;	99WO-US030720.
PR	30-DEC-1999;	99WO-US031243.
PR	30-DEC-1999;	99WO-US031274.
PR	05-JAN-2000;	2000WO-US000219.
PR	06-JAN-2000;	2000WO-US000277.
PR	06-JAN-2000;	2000WO-US000376.
PR	11-FEB-2000;	2000WO-US003565.
PR	18-FEB-2000;	2000WO-US004341.
PR	18-FEB-2000;	2000WO-US004342.
PR	22-FEB-2000;	2000WO-US004414.
PR	24-FEB-2000;	2000WO-US004914.
PR	24-FEB-2000;	2000WO-US005004.
PR	01-MAR-2000;	2000WO-US005601.
PR	02-MAR-2000;	2000WO-US005746.
PR	02-MAR-2000;	2000WO-US005841.
PR	10-MAR-2000;	2000WO-US006319.
PR	15-MAR-2000;	2000WO-US006884.
PR	20-MAR-2000;	2000WO-US007377.
PR	21-MAR-2000;	2000WO-US007532.
PR	30-MAR-2000;	2000WO-US008439.
PR	17-MAY-2000;	2000WO-US013705.
PR	22-MAY-2000;	2000WO-US014042.
PR	30-MAY-2000;	2000WO-US014941.
PR	02-JUN-2000;	2000WO-US015264.
PR	28-JUL-2000;	2000WO-US020710.
PR	11-AUG-2000;	2000WO-US022031.
PR	23-AUG-2000;	2000WO-US023522.
PR	24-AUG-2000;	2000WO-US023328.
PR	08-NOV-2000;	2000WO-US030952.
PR	10-NOV-2000;	2000WO-US030873.
PR	01-DEC-2000;	2000WO-US032678.
PR	20-DEC-2000;	2000US-00747259.
PR	20-DEC-2000;	2000WO-US034956.
PR	28-FEB-2001;	2001US-00796498.
PR	28-FEB-2001;	2001WO-US006520.
PR	01-MAR-2001;	2001WO-US006666.
PR	09-MAR-2001;	2001US-00802706.
PR	14-MAR-2001;	2001US-00808689.
PR	22-MAR-2001;	2001US-00816744.
PR	05-APR-2001;	2001US-00828366.
PR	10-MAY-2001;	2001US-00854208.
PR	10-MAY-2001;	2001US-00854280.
PR	18-MAY-2001;	2001US-00860216.
PR	25-MAY-2001;	2001US-00866028.
PR	25-MAY-2001;	2001US-00866034.
PR	25-MAY-2001;	2001WO-US017092.
PR	01-JUN-2001;	2001US-00872035.
PR	01-JUN-2001;	2001WO-US017800.
PR	05-JUN-2001;	2001US-00874503.
PR	14-JUN-2001;	2001US-00882636.
PR	19-JUN-2001;	2001US-00886342.
PR	20-JUN-2001;	2001WO-US019692.
PR	21-JUN-2001;	2001US-00887879.
PR	22-JUN-2001;	2001WO-US020116.
PR	29-JUN-2001;	2001WO-US021066.
PR	09-JUL-2001;	2001WO-US021735.
PR	18-JUL-2001;	2001US-00908827.
PR	06-AUG-2001;	2001US-00924419.
PR	09-AUG-2001;	2001US-00927796.
PR	16-AUG-2001;	2001US-00931836.

PR 14-SEP-1998; 98WO-US019177.
PR 16-SEP-1998; 98WO-US019330.
PR 17-SEP-1998; 98WO-US019437.
PR 07-OCT-1998; 98WO-US021141.
PR 29-OCT-1998; 98WO-US022991.
PR 29-OCT-1998; 98WO-US022992.
PR 20-NOV-1998; 98WO-US024855.
PR 01-DEC-1998; 98WO-US025108.
PR 05-JAN-1999; 99WO-US000106.
PR 08-MAR-1999; 99WO-US005028.
PR 10-MAR-1999; 99WO-US005190.
PR 20-APR-1999; 99WO-US008615.
PR 14-MAY-1999; 99WO-US010733.
PR 02-JUN-1999; 99WO-US012252.
PR 01-SEP-1999; 99WO-US020111.
PR 08-SEP-1999; 99WO-US020594.
PR 13-SEP-1999; 99WO-US020944.
PR 15-SEP-1999; 99WO-US021090.
PR 15-SEP-1999; 99WO-US021547.
PR 05-OCT-1999; 99WO-US023089.
PR 29-NOV-1999; 99WO-US028214.
PR 30-NOV-1999; 99WO-US028313.
PR 30-NOV-1999; 99WO-US028409.
PR 01-DEC-1999; 99WO-US028301.
PR 01-DEC-1999; 99WO-US028634.
PR 02-DEC-1999; 99WO-US028551.
PR 02-DEC-1999; 99WO-US028564.
PR 02-DEC-1999; 99WO-US028565.
PR 16-DEC-1999; 99WO-US030095.
PR 20-DEC-1999; 99WO-US030911.
PR 20-DEC-1999; 99WO-US030999.
PR 30-DEC-1999; 99WO-US031243.
PR 30-DEC-1999; 99WO-US031274.
PR 05-JAN-2000; 2000WO-US030219.
PR 06-JAN-2000; 2000WO-US030277.
PR 06-JAN-2000; 2000WO-US030376.
PR 11-FEB-2000; 2000WO-US030365.
PR 18-FEB-2000; 2000WO-US04341.
PR 18-FEB-2000; 2000WO-US04342.
PR 22-FEB-2000; 2000WO-US04414.
PR 24-FEB-2000; 2000WO-US04914.
PR 24-FEB-2000; 2000WO-US05004.
PR 01-MAR-2000; 2000WO-US05601.
PR 02-MAR-2000; 2000WO-US05746.
XX

(GETH) GENENTECH INC.

Baker KP, Beresini M, Deforge L, Desnoyers L, Filvaroff E, Gao W;
Gerritsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S;
Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;

WPI; 2003-352836/33.
N-PSDB; ACA67079.

New isolated PRO polypeptide useful for treating diabetes, rheumatoid
arthritis, sports injuries, obesity, hearing loss in mammals, stroke, or
heart attack.

Claim 12; Fig 172; 643pp; English.

The present invention relates to the isolation of novel human PRO
polypeptides, and the polynucleotide sequences encoding them. The PRO
polypeptides are secreted and transmembrane proteins. The PRO
polypeptides and polynucleotides are useful for preparing a medicament
useful in the treatment of diabetes, bone and/or cartilage disorders
(e.g. rheumatoid arthritis, sports injuries, osteoarthritis), obesity,
hyper- or hypo-insulinaemia, hearing loss, and coagulation disorders
(e.g. stroke, heart attack). Anti-PRO antibodies are useful in diagnostic
assays for PRO, by detecting its expression in specific cells, tissues or
serum, and for affinity purification of PRO from recombinant cell culture
or natural sources. ABU0870-ABU81144 represent the human PRO
polypeptides of the invention. Note: The sequence data for this patent
was obtained in electronic format directly from the USPTO web site at

CC seqdata.uspto.gov/psipsDIDentry.html
XX
SQ Sequence 182 AA;

Query Match 99.7%; Score 959; DB 6; Length 182;
Best Local Similarity 99.5%; Pred. No. 7.6e-95;
Matches 181; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEPQLGPEAAALRPGWLALLLWVSALSCSFSLPASSLSLVPQVRTSYNFGRTFLGLDKC 60
|||
Db 1 MEPQLGPEAAALRPGWLALLLWVSALSCSFSLPASSLSLVPQVRTSYNFGRTFLGLDKC 60

QY 61 NACIGTSICKKFFKEEIRSDNWLASHLGLPPDSLLSYPPANYSDDSKIWRPVEIFRLVSKY 120
|||
Db 61 NACIGTSICKKFFKEEIRSDNWLASHLGLPPDSLLSYPPANYSDDSKIWRPVEIFRLVSKY 120

QY 121 QNEISDRKICASASAPKTCISIERVLKTKTERFQKWLQAKRLTPDLVQDCHQGORELKFLCM 180
|||
Db 121 QNEISDRKICASASAPKTCISIERVLKTKTERFQKWLQAKRLTPDLVQDCHQGORELKFLCM 180

QY 181 LR 182
|||
Db 181 LR 182

RESULT 10
ABU66655
ID ABU66655 standard; protein; 182 AA.
XX
AC ABU66655;
XX
XX 23-MAY-2003 (first entry)
XX Human PRO polypeptide #86.
XX Human; PRO polypeptide; secreted and transmembrane protein;
KW tumour necrosis factor-alpha; TNF-alpha; blood; proliferation;
KW differentiation; chondrocyte; tumour; genetic disorder; cytostatic.
XX Homo sapiens.
XX
PN US2003036180-A1.
XX
PD 20-FEB-2003.
XX
XX 09-MAY-2002; 2002US-00143114.
XX
PR 31-MAR-1997; 97WO-US005230.
PR 12-JUN-1998; 98WO-US012456.
PR 14-JUL-1998; 98WO-US014552.
PR 28-AUG-1998; 98WO-US017888.
PR 10-SEP-1998; 98WO-US018824.
PR 14-SEP-1998; 98WO-US019093.
PR 14-SEP-1998; 98WO-US019094.
PR 14-SEP-1998; 98WO-US019177.
PR 16-SEP-1998; 98WO-US019330.
PR 17-SEP-1998; 98WO-US019437.
PR 07-OCT-1998; 98WO-US021141.
PR 29-OCT-1998; 98WO-US022991.
PR 29-OCT-1998; 98WO-US022992.
PR 20-NOV-1998; 98WO-US024855.
PR 01-DEC-1998; 98WO-US025108.
PR 05-JAN-1999; 99WO-US000106.
PR 08-MAR-1999; 99WO-US005028.
PR 10-MAR-1999; 99WO-US005190.
PR 20-APR-1999; 99WO-US008615.
PR 14-MAY-1999; 99WO-US010733.
PR 02-JUN-1999; 99WO-US012252.
PR 01-SEP-1999; 99WO-US020111.
PR 08-SEP-1999; 99WO-US020594.
PR 13-SEP-1999; 99WO-US020944.
PR 15-SEP-1999; 99WO-US021090.
PR 15-SEP-1999; 99WO-US021547.

PR 05-OCT-1999; 99WO-US023089.
PR 29-NOV-1999; 99WO-US028214.
PR 30-NOV-1999; 99WO-US028313.
PR 30-NOV-1999; 99WO-US028409.
PR 01-DEC-1999; 99WO-US028301.
PR 01-DEC-1999; 99WO-US028634.
PR 02-DEC-1999; 99WO-US028551.
PR 02-DEC-1999; 99WO-US028564.
PR 02-DEC-1999; 99WO-US028565.
PR 16-DEC-1999; 99WO-US030095.
PR 20-DEC-1999; 99WO-US030911.
PR 20-DEC-1999; 99WO-US030999.
PR 22-DEC-1999; 99WO-US030720.
PR 30-DEC-1999; 99WO-US031243.
PR 30-DEC-1999; 99WO-US031274.
PR 05-JAN-2000; 2000WO-US000219.
PR 06-JAN-2000; 2000WO-US000277.
PR 11-FEB-2000; 2000WO-US000376.
PR 18-FEB-2000; 2000WO-US003565.
PR 18-FEB-2000; 2000WO-US004341.
PR 22-FEB-2000; 2000WO-US004342.
PR 24-FEB-2000; 2000WO-US004414.
PR 24-FEB-2000; 2000WO-US004914.
PR 01-MAR-2000; 2000WO-US005004.
PR 02-MAR-2000; 2000WO-US005746.
PR 02-MAR-2000; 2000WO-US005841.
PR 10-MAR-2000; 2000WO-US006319.
PR 15-MAR-2000; 2000WO-US006884.
PR 20-MAR-2000; 2000WO-US007377.
PR 21-MAR-2000; 2000WO-US007532.
PR 30-MAR-2000; 2000WO-US008439.
PR 17-MAY-2000; 2000WO-US013705.
PR 22-MAY-2000; 2000WO-US014042.
PR 30-MAY-2000; 2000WO-US014941.
PR 02-JUN-2000; 2000WO-US015264.
PR 28-JUL-2000; 2000WO-US020710.
PR 11-AUG-2000; 2000WO-US022031.
PR 23-AUG-2000; 2000WO-US023522.
PR 24-AUG-2000; 2000WO-US023328.
PR 08-NOV-2000; 2000WO-US030952.
PR 10-NOV-2000; 2000WO-US030873.
PR 01-DEC-2000; 2000WO-US032678.
PR 20-DEC-2000; 2000US-00747259.
PR 20-DEC-2000; 2000WO-US034956.
PR 28-FEB-2001; 2001US-00796498.
PR 28-FEB-2001; 2001WO-US006520.
PR 01-MAR-2001; 2001WO-US006666.
PR 09-MAR-2001; 2001US-00802706.
PR 14-MAR-2001; 2001US-00808689.
PR 22-MAR-2001; 2001US-00816744.
PR 05-APR-2001; 2001US-00828366.
PR 10-MAY-2001; 2001US-00854208.
PR 10-MAY-2001; 2001US-00854280.
PR 18-MAY-2001; 2001US-00860216.
PR 25-MAY-2001; 2001US-00866028.
PR 25-MAY-2001; 2001US-00866034.
PR 25-MAY-2001; 2001WO-US017092.
PR 01-JUN-2001; 2001US-00872035.
PR 01-JUN-2001; 2001WO-US017800.
PR 05-JUN-2001; 2001US-00874503.
PR 14-JUN-2001; 2001US-00882636.
PR 19-JUN-2001; 2001US-00886342.
PR 20-JUN-2001; 2001WO-US019692.
PR 21-JUN-2001; 2001US-00887879.
PR 22-JUN-2001; 2001WO-US020116.
PR 29-JUN-2001; 2001WO-US021066.
PR 09-JUL-2001; 2001WO-US021735.
PR 18-JUL-2001; 2001US-00908827.
PR 06-AUG-2001; 2001US-00924419.
PR 09-AUG-2001; 2001US-00927796.
PR 16-AUG-2001; 2001US-00931836.
PR 19-DEC-2001; 2001US-00028072.

XX (GETH) GENENTECH INC.
PA Baker KP, Beresini M, Deforge L, Desnoyers L, Filvaroff E, Gao W;
XX Gerritsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S;
PI Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;
XX WPI; 2003-332040/31.
DR N-PSDB; ACA03688.
XX
PT New secreted and transmembrane PRO nucleic acids, useful for gene
PT therapy, in chromosome and gene mapping, as chromosome markers, in tissue
PT typing, and in chromosome identification.
XX
PS Claim 12; Fig 172; 660pp; English.
XX
CC The present invention relates to the isolation of novel human PRO
CC polypeptides, and the polynucleotide sequences encoding them. The PRO
CC polypeptides are secreted and transmembrane proteins. The PRO
CC polypeptides are useful for detecting other PRO polypeptides, for linking
CC bioactive molecules to cells expressing PRO polypeptides, for modulating
CC biological activities of cells expressing PRO polypeptides, and for
CC identifying agonists or antagonists. The PRO polypeptides are useful for
CC for stimulating the release of tumour necrosis factor (TNF)-alpha from
CC human blood, for stimulating the proliferation or differentiation of
CC chondrocytes, and detecting the presence of tumours. The polynucleotide
CC sequences encoding PRO polypeptides are useful as hybridisation probes,
CC in chromosome and gene mapping, in the generation of antisense RNA and
CC DNA, in the preparation of PRO polypeptides, for generating transgenic
CC animals or knockout animals, for the genetic analysis of individuals with
CC genetic disorders, and in gene therapy. ABU66570-ABU66844 represent the
CC human PRO polypeptides of the invention. Note: The sequence data for this
CC patent was obtained in electronic format directly from the USPTO web site
CC at seqdata.uspto.gov/psipdIDEntry.html
XX
SQ Sequence 182 AA;

Query Match 99.7%; Score 959; DB 6; Length 182;
Best Local Similarity 99.5%; Pred. No. 7.6e-95;
Matches 181; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEPQLGPEAAALRPGWLALLLWVSALSCSFSLSLPPDSSLVQVTSYNGRTGLDKC 60
|||
Db 1 MEPQLGPEAAALRPGWLALLLWVSALSCSFSLSLPPDSSLVQVTSYNGRTGLDKC 60

QY 61 NACIGTSICKKFFKEIRSDNWLASHGLPPDSSLVQVTSYNGRTGLDKC 120
|||
Db 61 NACIGTSICKKFFKEIRSDNWLASHGLPPDSSLVQVTSYNGRTGLDKC 120

QY 121 ONEISDRKICASAPKTCISIRVLRKTERFQKWLQAKRLTPDLVQDCHQGQRELKFLCM 180
|||
Db 121 ONEISDRKICASAPKTCISIRVLRKTERFQKWLQAKRLTPDLVQDCHQGQRELKFLCM 180

QY 181 LR 182
||
Db 181 LR 182

RESULT 11
ABU59736
ID ABU59736 standard; protein; 182 AA.
XX
AC ABU59736;
XX
DT 13-MAY-2003 (first entry)
XX
DE Novel secreted and transmembrane protein PRO3743.
XX
KW Human; PRO; hypertrophy of neonatal heart; angiogenesis; wound healing;
KW cardiac insufficiency disorder; cancer; tumour; immune response;
KW adrenal cortical capillary endothelial growth; c-fos induction;
KW vascular endothelial growth factor inhibition; VEGF inhibition;
KW endothelial cell growth inhibitor; T-lymphocytes stimulation;

retinal neurons cell survival; rod photoreceptor cell survival;
retinal disorder; retinitis pigmentosum; kidney disorder;
mammalian kidney mesangial cell proliferation; Berger disease;
dermatitis; herpeticiformis; Crohn's disease; chondrocyte proliferation;
ctondrocyte redifferentiation; sports injury; arthritis.

XW	OS	Homo sapiens.		PR	22-MAY-2000;	200WO-US014042.
XW	XX	US2003017563-A1.		PR	30-MAY-2000;	200WO-US014941.
XW	PN			PR	02-JUN-2000;	200WO-US015264.
XW	XX			PR	28-JUL-2000;	200WO-US020710.
KW	PD			PR	11-AUG-2000;	200WO-US022031.
XX	XX			PR	23-AUG-2000;	200WO-US023522.
OS	XX			PR	24-AUG-2000;	200WO-US023328.
XX	XX			PR	08-NOV-2000;	200WO-US030952.
PN	XX			PR	10-NOV-2000;	200WO-US030873.
XX	XX			PR	01-DEC-2000;	200WO-US032678.
PD	XX			PR	20-DEC-2000;	200OUS-00747259.
XX	XX			PR	20-DEC-2000;	200WO-US034956.
FF	XX			PR	28-FEB-2001;	2001US-00796498.
XX	XX			PR	01-MAR-2001;	2001WO-US006520.
PR	PR			PR	09-MAR-2001;	2001US-00802706.
PR	PR			PR	14-MAR-2001;	2001US-00808689.
PR	PR			PR	22-MAR-2001;	2001US-00816744.
PR	PR			PR	05-APR-2001;	2001US-00828366.
PR	PR			PR	10-MAY-2001;	2001US-00854208.
PR	PR			PR	10-MAY-2001;	2001US-00854280.
PR	PR			PR	18-MAY-2001;	2001US-00860216.
PR	PR			PR	25-MAY-2001;	2001US-00866028.
PR	PR			PR	25-MAY-2001;	2001US-00866034.
PR	PR			PR	25-MAY-2001;	2001WO-US017092.
PR	PR			PR	01-JUN-2001;	2001US-00872035.
PR	PR			PR	01-JUN-2001;	2001WO-US017800.
PR	PR			PR	05-JUN-2001;	2001US-00874503.
PR	PR			PR	14-JUN-2001;	2001US-00882636.
PR	PR			PR	19-JUN-2001;	2001US-00886342.
PR	PR			PR	20-JUN-2001;	2001WO-US019692.
PR	PR			PR	21-JUN-2001;	2001US-00887879.
PR	PR			PR	22-JUN-2001;	2001WO-US020116.
PR	PR			PR	29-JUN-2001;	2001WO-US021066.
PR	PR			PR	09-JUL-2001;	2001WO-US021735.
PR	PR			PR	18-JUL-2001;	2001US-00908827.
PR	PR			PR	06-AUG-2001;	2001US-00924419.
PR	PR			PR	09-AUG-2001;	2001US-00927796.
PR	PR			PR	16-AUG-2001;	2001US-00931836.
PR	XX			PR	19-DEC-2001;	2001US-00028072.
PA	XX			(GETH)	GENENTECH INC.	
PI	PI			Baker KP, Beresini M, Deforge L, Desnoyers L, Filvaroff E, Gao W;		
PI	PI			Gerritsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S;		
PI	PI			Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;		
XX	DR			WPI; 2003-148238/14.		
DR	DR			N-PSDB; ABX89226.		
XX	XX			Two hundred and seventy five nucleic acids encoding PRO polypeptides,		
PT	PT			useful for treating pericyte-associated tumors, diabetes and various bone		
PT	PT			and/or cartilage disorders, e.g. arthritis.		
XX	PS			Claim 12; Fig 172; 659pp; English.		
PS	XX			The invention describes an isolated human PRO polypeptide. The PRO		
XX	CC			polypeptides are useful in detecting PRO polypeptides in a sample, in		
CC	CC			linking a bioactive molecule to a cell expressing a PRO polypeptide, and		
CC	CC			in modulating at least one biological activity of a cell expressing a PRO		
CC	CC			polypeptide. PRO1312 stimulates hypertrophy of neonatal heart and is thus		
CC	CC			useful for treating cardiac insufficiency disorders. PRO1154 and PRO1186		
CC	CC			stimulate adrenal cortical capillary endothelial growth, and PRO536,		
CC	CC			PRO943, PRO829, PRO826, PRO1068 or PRO535, PRO826, PRO819, PRO1126,		
CC	CC			PRO1360 and PRO1387 induce c-fos in endothelial cells, and are thus		
CC	CC			useful for treating conditions or disorders where angiogenesis would be		
CC	CC			beneficial, e.g. wound healing and antagonist of this polypeptide are		
CC	CC			useful for treating cancerous tumours. PRO812 inhibits vascular		
CC	CC			endothelial growth factor (VEGF) stimulated proliferation of endothelial		
CC	CC			cells and is thus useful for inhibiting endothelial cell growth in		
CC	CC			mammals which would be beneficial in inhibiting tumour growth. PRO826,		
CC	CC			PRO1068, PRO1184, PRO1346 and PRO1375 stimulate proliferation of		
CC	CC			stimulated T-lymphocytes and are therapeutically useful for enhancing		

CC immune response. PRO828, PRO826, PRO1068 or PRO1132 enhance survival of
CC retinal neurons cells (PRO1132 is also enhances survival/proliferation of
CC rod photoreceptor cells) and therefore are useful for treating retinal
CC disorders of injuries, e.g. retinitis pigmentosum, AMD. PRO819, PRO813
CC and PRO11066 induce proliferation of mammalian kidney mesangial cells,
CC and therefore are useful for treating kidney disorders associated with
CC decreased mesangial cell function such as Berger disease or other
CC nephropathies associated with dermatitis, herpetiformis or Crohn's
CC disease. PRO1310, PRO844, PRO1312, PRO1192 and PRO1387 induce the
CC proliferation and/or redifferentiation of chondrocytes in culture and are
CC thus useful for treating sports injuries, and arthritis. This is the
CC amino acid sequence of a novel human PRO protein
XX
SQ Sequence 182 AA;

Query Match 99.7%; Score 959; DB 6; Length 182;
Best Local Similarity 99.5%; Pred. No. 7.6e-95;
Matches 181; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEPQLGPEAAALRPGWLALLLWVSALSCSFSLPASSLSLVPQVRTSYNFGRTFLGLDKC 60
Dd |||||
QY 61 NACIGTSICKKFFKBEIRSDNWLASHLGLPPDSLLSYPNYSDSKIWPRVEIFRLVSKY 120
Dd |||||
QY 121 QNEISDRKICASAPKTCISIERVLKTERFQKWLOAKRLTPDLVDCHQGQRELKFLCM 180
Dd |||||
QY 181 LR 182
Dd ||
181 LR 182

RESULT 12
ABO24926
ID ABO24926 standard; protein; 182 AA.
XX
AC ABO24926;
XX
DT 05-SEP-2003 (first entry)
DE Human secreted/transmembrane protein (PRO) #86.
XX
KW Human; PRO; secreted protein; transmembrane protein; tumour; cytostatic;
KW gene therapy; tumour necrosis factor-alpha; TNF-alpha; blood;
KW proteoglycan; cartilage; cytokine; peripheral blood mononuclear cell;
KW PBMC; glucose uptake; FFA; skeletal muscle cell; adipocyte cell;
KW chondrocyte cell proliferation; chondrocyte cell differentiation;
KW pericyte cell; inner ear utricular supporting cell; T-lymphocyte cell;
KW endothelial cell; A-peptide; factor VIIA.
XX
OS Homo sapiens.
XX
PN US2003036179-A1.
XX
PD 20-FEB-2003.
XX
PF 10-MAY-2002; 2002US-00142431.
XX
PR 31-MAR-1997; 97WO-US005230.
PR 12-JUN-1998; 98WO-US012456.
PR 14-JUL-1998; 98WO-US014552.
PR 28-AUG-1998; 98WO-US017888.
PR 10-SEP-1998; 98WO-US018824.
PR 14-SEP-1998; 98WO-US019093.
PR 14-SEP-1998; 98WO-US019094.
PR 14-SEP-1998; 98WO-US019177.
PR 16-SEP-1998; 98WO-US019330.
PR 17-SEP-1998; 98WO-US019437.
PR 07-OCT-1998; 98WO-US021141.
PR 29-OCT-1998; 98WO-US022991.
PR 29-OCT-1998; 98WO-US022992.
PR 20-NOV-1998; 98WO-US024855.
PR 01-DEC-1998; 98WO-US025108.
PR 05-JAN-1999; 99WO-US000106.
PR 08-MAR-1999; 99WO-US005028.
PR 10-MAR-1999; 99WO-US005190.
PR 20-APR-1999; 99WO-US008615.
PR 14-MAY-1999; 99WO-US010733.
PR 02-JUN-1999; 99WO-US012252.
PR 01-SEP-1999; 99WO-US020111.
PR 08-SEP-1999; 99WO-US020594.
PR 13-SEP-1999; 99WO-US020944.
PR 15-SEP-1999; 99WO-US021090.
PR 15-SEP-1999; 99WO-US021547.
PR 05-OCT-1999; 99WO-US023089.
PR 29-NOV-1999; 99WO-US028214.
PR 30-NOV-1999; 99WO-US028313.
PR 30-NOV-1999; 99WO-US028409.
PR 01-DEC-1999; 99WO-US028301.
PR 01-DEC-1999; 99WO-US028634.
PR 02-DEC-1999; 99WO-US028551.
PR 02-DEC-1999; 99WO-US028564.
PR 02-DEC-1999; 99WO-US028565.
PR 16-DEC-1999; 99WO-US030095.
PR 20-DEC-1999; 99WO-US030911.
PR 20-DEC-1999; 99WO-US030999.
PR 22-DEC-1999; 99WO-US030720.
PR 30-DEC-1999; 99WO-US031243.
PR 30-DEC-1999; 99WO-US031274.
PR 05-JAN-2000; 2000WO-US000219.
PR 06-JAN-2000; 2000WO-US000277.
PR 06-JAN-2000; 2000WO-US000376.
PR 11-FEB-2000; 2000WO-US003565.
PR 18-FEB-2000; 2000WO-US004341.
PR 18-FEB-2000; 2000WO-US004342.
PR 22-FEB-2000; 2000WO-US004414.
PR 24-FEB-2000; 2000WO-US004914.
PR 24-FEB-2000; 2000WO-US005004.
PR 01-MAR-2000; 2000WO-US005601.
PR 02-MAR-2000; 2000WO-US005746.
PR 02-MAR-2000; 2000WO-US005841.
PR 10-MAR-2000; 2000WO-US006319.
PR 15-MAR-2000; 2000WO-US006884.
PR 20-MAR-2000; 2000WO-US007377.
PR 21-MAR-2000; 2000WO-US007532.
PR 30-MAR-2000; 2000WO-US008439.
PR 17-MAY-2000; 2000WO-US013705.
PR 22-MAY-2000; 2000WO-US014042.
PR 30-MAY-2000; 2000WO-US014941.
PR 02-JUN-2000; 2000WO-US015264.
PR 28-JUL-2000; 2000WO-US020710.
PR 11-AUG-2000; 2000WO-US022031.
PR 23-AUG-2000; 2000WO-US023522.
PR 24-AUG-2000; 2000WO-US023328.
PR 08-NOV-2000; 2000WO-US030952.
PR 10-NOV-2000; 2000WO-US030873.
PR 01-DEC-2000; 2000WO-US032678.
PR 20-DEC-2000; 2000US-00747259.
PR 20-DEC-2000; 2000WO-US034956.
PR 28-FEB-2001; 2001US-00796498.
PR 28-FEB-2001; 2001WO-US006520.
PR 01-MAR-2001; 2001WO-US006666.
PR 09-MAR-2001; 2001US-00802706.
PR 14-MAR-2001; 2001US-00808689.
PR 22-MAR-2001; 2001US-00816744.
PR 05-APR-2001; 2001US-00828366.
PR 10-MAY-2001; 2001US-00854208.
PR 10-MAY-2001; 2001US-00854280.
PR 18-MAY-2001; 2001US-00860216.
PR 25-MAY-2001; 2001US-00866028.
PR 25-MAY-2001; 2001US-00866034.
PR 25-MAY-2001; 2001WO-US017092.

PR 01-JUN-2001; 2001US-00872035.
PR 01-JUN-2001; 2001WO-US017800.
PR 05-JUN-2001; 2001US-00874503.
PR 14-JUN-2001; 2001US-00882636.
PR 19-JUN-2001; 2001US-00886342.
PR 20-JUN-2001; 2001WO-US019692.
PR 21-JUN-2001; 2001US-00887879.
PR 22-JUN-2001; 2001WO-US020116.
PR 29-JUN-2001; 2001WO-US021066.
PR 09-JUL-2001; 2001WO-US021735.
PR 18-JUL-2001; 2001US-00908827.
PR 06-AUG-2001; 2001US-00924419.
PR 09-AUG-2001; 2001US-00927796.
PR 16-AUG-2001; 2001US-00931836.
PR 19-DEC-2001; 2001US-00028072.
XX
PA (GETH) GENENTECH INC.
XX
PI Baker KP, Beresini M, Deforge L, Desnoyers L, Filvaroff E, Gao W;
PI Gerritsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S;
PI Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;
XX
DR WPI; 2003-466355/44.
DR N-PSDB; ACD41880.
XX
PT New isolated nucleic acid encoding a PRO polypeptide, e.g. PRO1114 or
PT PRO4978, useful in molecular biology, chromosome and gene mapping, in
PT generating antisense RNA and DNA, and in gene therapy.
XX
PS Claim 12; Fig 172; 659pp; English.
XX
CC The invention relates to an isolated nucleic acid comprising at least 80%
CC sequence identity to a PRO (secreted and transmembrane protein) cDNA
CC comprising a nucleic acid (a) encoding a PRO polypeptide, or its
CC extracellular domain (with or without its associated signal peptide),
CC which comprises any of the 275 120-850 residue amino acid sequences,
CC given in the specification; (b) comprising any of the 275 300-3500
CC nucleotide sequences, given in the specification; or (c) comprising the
CC full-length coding sequence of the nucleotide sequences given in the
CC specification, or of the DNA deposited under any of the American Type
CC Culture Collection (ATCC) Accession Numbers listed in the specification.
CC Also included are a vector comprising the novel nucleic acid, a host cell
CC comprising the vector, producing a PRO polypeptide, the isolated PRO
CC polypeptides detailed above, a chimaeric molecule comprising the PRO
CC polypeptide of fused to a heterologous amino acid sequence, an anti-PRO
CC antibody, detecting a PRO polypeptide in a sample suspected of containing
CC the PRO polypeptide, linking a bioactive molecule to a cell expressing a
CC PRO polypeptide, modulating at least one biological activity of a cell
CC expressing a PRO polypeptide, stimulating the release of tumour necrosis
CC factor-alpha (TNF-alpha) from human blood, (or proteoglycans from
CC cartilage or cytokine from peripheral blood mononuclear cells (PBMC)),
CC modulating the uptake of glucose or FFA by skeletal muscle cells or
CC adipocyte cells, stimulating the proliferation or differentiation of
CC chondrocyte cells (or proliferation of or gene expression in pericyte
CC cells), stimulating the proliferation of inner ear utricular supporting
CC cells (or of T-lymphocyte cells, or of endothelial cells), inhibiting the
CC binding of A-peptide to factor VIIA, or differentiation of adipocyte
CC cells, detecting the presence of a tumour in a mammal and a
CC oligonucleotide probe derived from any of the nucleotide sequences given
CC in the specification. The polynucleotide is useful in molecular biology,
CC including uses as hybridisation probes, in chromosome and gene mapping,
CC in generating antisense RNA and DNA, and in gene therapy. The
CC polynucleotide may also be used in preparing PRO polypeptides by
CC recombinant techniques, and in generating either transgenic animals or
CC knock-out animals which, in turn, are useful in the development and
CC screening of therapeutically useful reagents. The PRO polypeptide or the
CC antibody is used in preparing a medicament for treating a condition
CC responsive to the polypeptide or antibody, such as tumours, and in
CC various diagnostic assays. The present sequence represents a PRO
CC polypeptide
XX
SQ Sequence 182 AA;

Query Match 99.7%; Score 959; DB 6; Length 182;
Best Local Similarity 99.5%; Pred. No. 7.6e-95;
Matches 181; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 MEPQLGPEAAALRPGWLALLLWVSALSCSFSLSLPASSLSLVLPQVTSYNFGRTELGLDKC 60
|||
Db 1 MEPQLGPEAAALRPGWLALLLWVSALSCSFSLSLPASSLSLVLPQVTSYNFGRTELGLDKC 60
QY 61 NACIGTSICKKFFKEEIRSDNNWLASHLGLPPDSLLSYPNYSDDSKIWRPVEIFRLVSKY 120
|||
Db 61 NACIGTSICKKFFKEEIRSDNNWLASHLGLPPDSLLSYPNYSDDSKIWRPVEIFRLVSKY 120
QY 121 QNEISDRKICASASAPKTCISIERVLKTERFQKWLQAKRLTPDLVQDCHQGQRELKFLCM 180
|||
Db 121 QNEISDRKICASASAPKTCISIERVLKTERFQKWLQAKRLTPDLVQDCHQGQRELKFLCM 180
QY 181 LR 182
||
Db 181 LR 182
RESULT 13
ABU66931
ID ABU66931 standard; protein; 182 AA.
XX
AC ABU66931;
XX
DT 27-MAY-2003 (first entry)
XX
DE Human secreted/transmembrane, PRO, protein SEQ ID 172.
XX
KW Human; secreted protein; transmembrane protein; PRO;
KW inflammatory disease; organ failure; atherosclerosis; cardiac injury;
KW infertility; birth defects; premature aging; AIDS; biosensor;
KW acquired immunodeficiency syndrome; cancer; diabetic complication;
KW bioreactor; tumour.
XX
OS Homo sapiens.
XX
PN US2003032155-A1.
XX
PD 13-FEB-2003.
XX
PF 03-MAY-2002; 2002US-00137865.
XX
PR 31-MAR-1997; 97WO-US005230.
PR 12-JUN-1998; 98WO-US012456.
PR 14-JUL-1998; 98WO-US014552.
PR 28-AUG-1998; 98WO-US017888.
PR 10-SEP-1998; 98WO-US018824.
PR 14-SEP-1998; 98WO-US019093.
PR 14-SEP-1998; 98WO-US019094.
PR 14-SEP-1998; 98WO-US019177.
PR 16-SEP-1998; 98WO-US019330.
PR 17-SEP-1998; 98WO-US019437.
PR 07-OCT-1998; 98WO-US021141.
PR 29-OCT-1998; 98WO-US022991.
PR 29-OCT-1998; 98WO-US022992.
PR 20-NOV-1998; 98WO-US024855.
PR 01-DEC-1998; 98WO-US025108.
PR 05-JAN-1999; 99WO-US000106.
PR 08-MAR-1999; 99WO-US005028.
PR 10-MAR-1999; 99WO-US005190.
PR 20-APR-1999; 99WO-US008615.
PR 14-MAY-1999; 99WO-US010733.
PR 02-JUN-1999; 99WO-US012252.
PR 01-SEP-1999; 99WO-US020111.
PR 08-SEP-1999; 99WO-US020594.
PR 13-SEP-1999; 99WO-US020944.
PR 15-SEP-1999; 99WO-US021090.
PR 05-OCT-1999; 99WO-US021547.
PR 29-NOV-1999; 99WO-US028214.

PR 30-NOV-1999; 99WO-US028313.
PR 30-NOV-1999; 99WO-US028409.
PR 01-DEC-1999; 99WO-US028301.
PR 01-DEC-1999; 99WO-US028634.
PR 02-DEC-1999; 99WO-US028551.
PR 02-DEC-1999; 99WO-US028564.
PR 02-DEC-1999; 99WO-US028565.
PR 16-DEC-1999; 99WO-US030095.
PR 20-DEC-1999; 99WO-US030911.
PR 20-DEC-1999; 99WO-US030999.
PR 22-DEC-1999; 99WO-US030720.
PR 30-DEC-1999; 99WO-US031243.
PR 30-DEC-1999; 99WO-US031274.
PR 05-JAN-2000; 2000WO-US000219.
PR 06-JAN-2000; 2000WO-US000277.
PR 06-JAN-2000; 2000WO-US000376.
PR 11-FEB-2000; 2000WO-US003565.
PR 18-FEB-2000; 2000WO-US004341.
PR 18-FEB-2000; 2000WO-US004342.
PR 22-FEB-2000; 2000WO-US004414.
PR 24-FEB-2000; 2000WO-US004914.
PR 24-FEB-2000; 2000WO-US005004.
PR 01-MAR-2000; 2000WO-US005601.
PR 02-MAR-2000; 2000WO-US005746.
PR 02-MAR-2000; 2000WO-US005841.
PR 10-MAR-2000; 2000WO-US006319.
PR 15-MAR-2000; 2000WO-US006884.
PR 20-MAR-2000; 2000WO-US007377.
PR 21-MAR-2000; 2000WO-US007532.
PR 30-MAR-2000; 2000WO-US008439.
PR 17-MAY-2000; 2000WO-US013705.
PR 22-MAY-2000; 2000WO-US014042.
PR 30-MAY-2000; 2000WO-US014941.
PR 02-JUN-2000; 2000WO-US015264.
PR 28-JUL-2000; 2000WO-US020710.
PR 11-AUG-2000; 2000WO-US022031.
PR 23-AUG-2000; 2000WO-US023522.
PR 24-AUG-2000; 2000WO-US023328.
PR 08-NOV-2000; 2000WO-US030952.
PR 10-NOV-2000; 2000WO-US030873.
PR 01-DEC-2000; 2000WO-US032678.
PR 20-DEC-2000; 2000US-00747259.
PR 20-DEC-2000; 2000WO-US034956.
PR 28-FEB-2001; 2001US-00796498.
PR 28-FEB-2001; 2001WO-US006520.
PR 01-MAR-2001; 2001WO-US006666.
PR 09-MAR-2001; 2001US-00802706.
PR 14-MAR-2001; 2001US-00808689.
PR 22-MAR-2001; 2001US-00816744.
PR 05-APR-2001; 2001US-00828366.
PR 10-MAY-2001; 2001US-00854208.
PR 10-MAY-2001; 2001US-00854280.
PR 18-MAY-2001; 2001US-00860216.
PR 25-MAY-2001; 2001US-00866028.
PR 25-MAY-2001; 2001US-00866034.
PR 25-MAY-2001; 2001WO-US017092.
PR 01-JUN-2001; 2001US-00872035.
PR 01-JUN-2001; 2001WO-US017800.
PR 05-JUN-2001; 2001US-00874503.
PR 14-JUN-2001; 2001US-00882636.
PR 19-JUN-2001; 2001US-00886342.
PR 20-JUN-2001; 2001WO-US019692.
PR 21-JUN-2001; 2001US-00887879.
PR 22-JUN-2001; 2001WO-US020116.
PR 29-JUN-2001; 2001WO-US021066.
PR 09-JUL-2001; 2001WO-US021735.
PR 18-JUL-2001; 2001US-00908827.
PR 06-AUG-2001; 2001US-00924419.
PR 09-AUG-2001; 2001US-00927796.
PR 16-AUG-2001; 2001US-00931836.
PR 19-DEC-2001; 2001US-00028072.

XX (GETH) GENENTECH INC. PA

XX Baker KP, Beresini M, Deforge L, Desnoyers L, Filvaroff E, Gao W;
PI Gerritsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S;
PI Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;
XX WPI; 2003-331925/31.
DR N-PSDB; ACA04109.
XX
PT New secreted and transmembrane nucleic acids and polypeptides, designated
as PRO, useful for treating inflammation, organ failure, atherosclerosis,
PT cardiac injury, infertility, birth defects, premature aging, AIDS, or
PT cancer.
XX
PS Claim 12; Fig 172; 659pp; English.
XX
CC The invention relates to an isolated nucleic acid comprising, or which is
at least 80% identical to, or the full-length coding sequence of, any of
CC the 275 nucleotide sequences, encoding the corresponding PRO polypeptide
CC (one of 275 secreted or transmembrane proteins). The nucleic acid further
CC comprises the full-length coding sequence of the DNA deposited under
CC American Type Culture Collection (ATCC) accession number in a list given
CC in the specification. Also included are vectors and host cells for
CC producing PRO proteins, PRO fusion proteins, anti-PRO antibodies, PRO
CC extracellular domains and mature sequences, methods of detecting PRO
CC proteins, methods for stimulating the release of TNF-alpha (tumour
CC necrosis factor alpha) from human blood, (and the proliferation of
CC differentiation of chondrocyte cells, the proliferation of, or gene
CC expression in pericyte cells, the release of proteoglycans from
CC cartilage, proliferation of inner ear utricular supporting cells, the
CC proliferation of T-lymphocyte cells, the release of a cytokine from
CC peripheral blood mononuclear cells (PBMC), or the proliferation of
CC endothelial cells), a method for modulating the uptake of glucose or free
CC fatty acid (FFA) by skeletal muscle cells, a method for inhibiting the
CC binding of A-peptide to factor VIIA, or the differentiation of adipocyte
CC cells, a method for detecting the presence of a tumour in a mammal and an
CC oligonucleotide probe derived from any of the nucleotide sequences cited
CC above. The nucleic acids and polypeptides are useful for treating
CC inflammatory diseases, organ failure, atherosclerosis, cardiac injury,
CC infertility, birth defects, premature aging, AIDS (acquired
CC immunodeficiency syndrome), cancer, or diabetic complications. The
CC nucleic acids are useful as hybridisation probes, in chromosome and gene
CC mapping, and in generating antisense RNA or DNA. The polypeptides are
CC useful as pharmaceuticals, diagnostics, biosensors or bioreactors. Both
CC are useful in tissue typing. The present sequence represents a PRO
CC protein of the invention
XX
SQ Sequence 182 AA;

Query Match 99.7%; Score 359; DB 6; Length 182;
Best Local Similarity 99.5%; Pred. No. 7.6e-95;
Matches 181; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEPQLGPEAAALRPGWLALLLWVLSALSCFSFLPASSLSSLPQVRTSYNFGRTFLGLDKC 60
|||
DB 1 MEPQLGPEAAALRPGWLALLLWVLSALSCFSFLPASSLSSLPQVRTSYNFGRTFLGLDKC 60

QY 61 NACIGTSICKKFFKBEIRSDNWLASHLGLPPDLSLLSYPNYSDDSKIWRPVEIFRLVSKY 120
|||
DB 61 NACIGTSICKKFFKBEIRSDNWLASHLGLPPDLSLLSYPNYSDDSKIWRPVEIFRLVSKY 120

QY 121 ONEISDRKICASAPKTCISIERVLRKTERFQKWLQAKRLTPDLVQDCHQGRELKFLCM 180
|||
DB 121 ONEISDRKICASAPKTCISIERVLRKTERFQKWLQAKRLTPDLVQDCHQGRELKFLCM 180

QY 181 LR 182
|||
DB 181 LR 182

RESULT 14
ADA45691
ID ADA45691 standard; protein; 182 AA.
XX

AC ADA45691;
XX
DT 20-NOV-2003 (first entry)
XX
DE Novel human secreted and transmembrane protein PRO3743.
XX
KW Human; secreted and transmembrane protein; PRO;
KW Tumour necrosis factor alpha release; TNF-alpha release;
KW glucose uptake modulator; FFA uptake modulator;
KW cell proliferation stimulator; cell differentiation stimulator;
KW cell differentiation inhibitor; cytokine release stimulator; tumour;
KW lung tumour; colon tumour; breast tumour; prostate tumour; rectal tumour;
KW cervical tumour; liver tumour; chromosome mapping; gene mapping;
KW gene therapy; chromosome identification; chromosome marker.
XX
OS Homo sapiens.
XX
PN US2003022328-A1.
XX
PD 30-JAN-2003.
XX
PF 16-APR-2002; 2002US-00123904.
XX
PR 31-MAR-1997; 97WO-US005230.
PR 12-JUN-1998; 98WO-US012456.
PR 14-JUL-1998; 98WO-US014552.
PR 28-AUG-1998; 98WO-US017888.
PR 10-SEP-1998; 98WO-US018824.
PR 14-SEP-1998; 98WO-US019093.
PR 14-SEP-1998; 98WO-US019094.
PR 14-SEP-1998; 98WO-US019177.
PR 16-SEP-1998; 98WO-US019330.
PR 17-SEP-1998; 98WO-US019437.
PR 07-OCT-1998; 98WO-US021141.
PR 29-OCT-1998; 98WO-US022991.
PR 29-OCT-1998; 98WO-US022992.
PR 20-NOV-1998; 98WO-US024855.
PR 01-DEC-1998; 98WO-US025108.
PR 05-JAN-1999; 99WO-US000106.
PR 08-MAR-1999; 99WO-US005028.
PR 10-MAR-1999; 99WO-US005190.
PR 20-APR-1999; 99WO-US008615.
PR 14-MAY-1999; 99WO-US010733.
PR 02-JUN-1999; 99WO-US012252.
PR 01-SEP-1999; 99WO-US020111.
PR 08-SEP-1999; 99WO-US020594.
PR 13-SEP-1999; 99WO-US020944.
PR 15-SEP-1999; 99WO-US021090.
PR 15-SEP-1999; 99WO-US021547.
PR 05-OCT-1999; 99WO-US023089.
PR 29-NOV-1999; 99WO-US028214.
PR 30-NOV-1999; 99WO-US028313.
PR 30-NOV-1999; 99WO-US028409.
PR 01-DEC-1999; 99WO-US028301.
PR 01-DEC-1999; 99WO-US028634.
PR 02-DEC-1999; 99WO-US028551.
PR 02-DEC-1999; 99WO-US028564.
PR 02-DEC-1999; 99WO-US028565.
PR 16-DEC-1999; 99WO-US030095.
PR 20-DEC-1999; 99WO-US030911.
PR 20-DEC-1999; 99WO-US030999.
PR 22-DEC-1999; 99WO-US030720.
PR 30-DEC-1999; 99WO-US031243.
PR 30-DEC-1999; 99WO-US031274.
PR 05-JAN-2000; 2000WO-US000219.
PR 06-JAN-2000; 2000WO-US000277.
PR 06-JAN-2000; 2000WO-US000376.
PR 11-FEB-2000; 2000WO-US003565.
PR 18-FEB-2000; 2000WO-US004341.
PR 18-FEB-2000; 2000WO-US004342.
PR 22-FEB-2000; 2000WO-US004414.
PR 24-FEB-2000; 2000WO-US004914.
PR 24-FEB-2000; 2000WO-US005004.
PR 24-FEB-2000; 2000WO-US005004.

PR 01-MAR-2000; 2000WO-US005601.
PR 02-MAR-2000; 2000WO-US005746.
PR 02-MAR-2000; 2000WO-US005841.
PR 10-MAR-2000; 2000WO-US006319.
PR 15-MAR-2000; 2000WO-US006884.
PR 20-MAR-2000; 2000WO-US007377.
PR 21-MAR-2000; 2000WO-US007532.
PR 30-MAR-2000; 2000WO-US008439.
PR 17-MAY-2000; 2000WO-US013705.
PR 22-MAY-2000; 2000WO-US014042.
PR 30-MAY-2000; 2000WO-US014941.
PR 02-JUN-2000; 2000WO-US015264.
PR 28-JUL-2000; 2000WO-US020710.
PR 11-AUG-2000; 2000WO-US022031.
PR 23-AUG-2000; 2000WO-US023522.
PR 24-AUG-2000; 2000WO-US023328.
PR 08-NOV-2000; 2000WO-US030952.
PR 10-NOV-2000; 2000WO-US030873.
PR 01-DEC-2000; 2000WO-US032678.
PR 20-DEC-2000; 2000US-00747259.
PR 20-DEC-2000; 2000WO-US034956.
PR 28-FEB-2001; 2001US-00796498.
PR 28-FEB-2001; 2001WO-US006520.
PR 01-MAR-2001; 2001WO-US006666.
PR 09-MAR-2001; 2001US-00802706.
PR 14-MAR-2001; 2001US-00808689.
PR 22-MAR-2001; 2001US-00816744.
PR 05-APR-2001; 2001US-00828366.
PR 10-MAY-2001; 2001US-00854208.
PR 10-MAY-2001; 2001US-00854280.
PR 18-MAY-2001; 2001US-00860216.
PR 25-MAY-2001; 2001US-00866028.
PR 25-MAY-2001; 2001US-00866034.
PR 25-MAY-2001; 2001WO-US017092.
PR 01-JUN-2001; 2001US-00872035.
PR 01-JUN-2001; 2001WO-US017800.
PR 05-JUN-2001; 2001US-00874503.
PR 14-JUN-2001; 2001US-00882636.
PR 19-JUN-2001; 2001US-00886342.
PR 20-JUN-2001; 2001WO-US019692.
PR 21-JUN-2001; 2001US-00887879.
PR 22-JUN-2001; 2001WO-US020116.
PR 29-JUN-2001; 2001WO-US021066.
PR 09-JUL-2001; 2001WO-US021735.
PR 18-JUL-2001; 2001US-00908827.
PR 06-AUG-2001; 2001US-00924419.
PR 09-AUG-2001; 2001US-00927796.
PR 16-AUG-2001; 2001US-00931836.
PR 19-DEC-2001; 2001US-00028072.
XX
PA (GETH) GENENTECH INC.
XX
PI Baker KP, Beresini M, DeForge L, Desnoyers L, Filvaroff E, Gao W;
PI Gerritsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S;
PI Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;
XX
DR WPI; 2003-584997/55.
DR N-PSDB; ADA45690.
XX
PT Novel secreted and transmembrane polypeptide for modulating biological
PT activity of cell expressing the polypeptide, identifying agonists or
PT antagonists of polypeptide, and as molecular weight markers.
XX
PS Claim 12; Fig 172; 659pp; English.
XX
CC The invention describes 305 nucleic acids encoding PRO (secreted and
CC transmembrane) polypeptides (I). (I) is useful for stimulating the
CC release of TNF-alpha from human blood, for modulating the uptake of
CC glucose or FFA by skeletal muscle cells or adipocyte cells, for
CC stimulating the proliferation or differentiation of chondrocyte cells,
CC for stimulating the proliferation of or gene expression in pericyte
CC cells, for stimulating the release of proteoglycans from cartilage, for
CC stimulating the proliferation of inner ear utricular supporting cells,

CC for stimulating the proliferation of T-lymphocyte cells, for stimulating
CC the release of a cytokine from PBMC cells, for inhibiting the binding of
CC A-peptide to factor VIIA, for inhibiting the differentiation of adipocyte
CC cells, for stimulating proliferation of endothelial cells, for detecting
CC the presence of tumour in a mammal. The tumour is lung, colon, breast,
CC prostate, rectal, cervical or liver tumour. The oligonucleotide probes
CC are useful for isolating genomic and cDNA nucleotide sequences or
CC antisense probes. (I) is also useful as therapeutic agent. PRO is useful
CC in assays to identify other proteins or molecules involved in binding
CC interaction. A polynucleotide (II) encoding (I) is useful in chromosome
CC and gene mapping, in generation of antisense RNA and DNA, in the
CC preparation of PRO polypeptide, for generating transgenic animals or
CC knockout animals which in turn are useful in the development and
CC screening of therapeutically useful reagents, in gene therapy, for
CC chromosome identification, as chromosome marker, and for generating
CC probes. An anti-(I)-antibody is useful in diagnostic assays for PRO, e.g.
CC detecting its expression in specific cells, tissues or serum, and for
CC affinity purification of PRO from recombinant cell culture or natural
CC sources. (I) and (II) are useful for tissue typing. This is the amino
CC acid sequence of a novel human secreted and transmembrane PRO
CC polypeptide.

SQ Sequence 182 AA;

Query Match 99.7%; Score 959; DB 6; Length 182;
Best Local Similarity 99.5%; Pred. No. 7.6e-95;
Matches 181; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEPQLGPEAAALRPGWLALLLWVSALSCSFSLPASSLSPQVRTSYNFGRTFLGLDKC 60
Db |||||
1 MEPQLGPEAAALRPGWLALLLWVSALSCSFSLPASSLSPQVRTSYNFGRTFLGLDKC 60

QY 61 NACIGTSICKFFKEEIRSDNWLASHLGLPDSLLSYPANYSDDSKIWRPVEIFRLVSKY 120
Db |||||
61 NACIGTSICKFFKEEIRSDNWLASHLGLPDSLLSYPANYSDDSKIWRPVEIFRLVSKY 120

QY 121 QNEISDRKICASAPKTCISIERVLKRTKTERFQKWLQAKRLTPDLVDCHQGQRELKFLCM 180
Db |||||:|||||
121 QNEISDRKICASAPKTCISIERVLKRTKTERFQKWLQAKRLTPDLVDCHQGQRELKFLCM 180

QY 181 LR 182
Db ||
181 LR 182

RESULT 15

ADA76122
ID ADA76122 standard; protein; 182 AA.

AC ADA76122;

DT 20-NOV-2003 (first entry)

DE Human PRO polypeptide #86.

XX
KW Human; PRO; secreted polypeptide; transmembrane polypeptide;
KW tumour necrosis factor-alpha; TNF-alpha; chondrocyte cell; tumour;
KW cancer; adrenal; lung; colon; breast; prostate; rectum; kidney; cervix;
KW liver; microvascular endothelial cell; glucose; FFA;
KW skeletal muscle cell; adipocyte cell; pericyte cell;
KW inner ear utricular supporting cell; T-lymphocyte cell;
KW endothelial cell tube formation; bone disorder; cartilage disorder;
KW sports injury; proteoglycan; articular cartilage defect; osteoarthritis;
KW rheumatoid arthritis; haemoglobin-associated disorder thalassaemia;
KW immune system cell infiltration.

XX Homo sapiens.

OS US2003073212-A.

XX

PD 17-APR-2003.

XX 16-APR-2002; 2002US-00123903.

XX
PR 31-MAR-1997; 97WO-US005230.
PR 12-JUN-1998; 98WO-US012456.
PR 14-JUL-1998; 98WO-US014552.
PR 28-AUG-1998; 98WO-US017888.
PR 10-SEP-1998; 98WO-US018824.
PR 14-SEP-1998; 98WO-US019093.
PR 14-SEP-1998; 98WO-US019094.
PR 14-SEP-1998; 98WO-US019177.
PR 16-SEP-1998; 98WO-US019330.
PR 17-SEP-1998; 98WO-US019437.
PR 07-OCT-1998; 98WO-US021141.
PR 29-OCT-1998; 98WO-US022991.
PR 29-OCT-1998; 98WO-US022992.
PR 20-NOV-1998; 98WO-US024855.
PR 01-DEC-1998; 98WO-US025108.
PR 05-JAN-1999; 99WO-US000106.
PR 08-MAR-1999; 99WO-US005028.
PR 10-MAR-1999; 99WO-US005190.
PR 20-APR-1999; 99WO-US008615.
PR 14-MAY-1999; 99WO-US010733.
PR 02-JUN-1999; 99WO-US012252.
PR 01-SEP-1999; 99WO-US020111.
PR 08-SEP-1999; 99WO-US020594.
PR 13-SEP-1999; 99WO-US020944.
PR 15-SEP-1999; 99WO-US021090.
PR 15-SEP-1999; 99WO-US021547.
PR 05-OCT-1999; 99WO-US023089.
PR 29-NOV-1999; 99WO-US028214.
PR 30-NOV-1999; 99WO-US028313.
PR 30-NOV-1999; 99WO-US028409.
PR 01-DEC-1999; 99WO-US028301.
PR 01-DEC-1999; 99WO-US028634.
PR 02-DEC-1999; 99WO-US028551.
PR 02-DEC-1999; 99WO-US028564.
PR 02-DEC-1999; 99WO-US028565.
PR 16-DEC-1999; 99WO-US030095.
PR 20-DEC-1999; 99WO-US030911.
PR 20-DEC-1999; 99WO-US030999.
PR 22-DEC-1999; 99WO-US030720.
PR 30-DEC-1999; 99WO-US031243.
PR 30-DEC-1999; 99WO-US031274.
PR 05-JAN-2000; 2000WO-US000219.
PR 06-JAN-2000; 2000WO-US000277.
PR 06-JAN-2000; 2000WO-US000376.
PR 11-FEB-2000; 2000WO-US003565.
PR 18-FEB-2000; 2000WO-US004341.
PR 18-FEB-2000; 2000WO-US004342.
PR 22-FEB-2000; 2000WO-US004414.
PR 24-FEB-2000; 2000WO-US004914.
PR 24-FEB-2000; 2000WO-US005004.
PR 01-MAR-2000; 2000WO-US005601.
PR 02-MAR-2000; 2000WO-US005746.
PR 02-MAR-2000; 2000WO-US005841.
PR 13-MAR-2000; 2000WO-US006319.
PR 15-MAR-2000; 2000WO-US006884.
PR 20-MAR-2000; 2000WO-US007377.
PR 21-MAR-2000; 2000WO-US007532.
PR 30-MAR-2000; 2000WO-US008439.
PR 17-MAY-2000; 2000WO-US013705.
PR 22-MAY-2000; 2000WO-US014042.
PR 30-MAY-2000; 2000WO-US014941.
PR 02-JUN-2000; 2000WO-US015264.
PR 28-JUL-2000; 2000WO-US020710.
PR 11-AUG-2000; 2000WO-US022031.
PR 23-AUG-2000; 2000WO-US023522.
PR 24-AUG-2000; 2000WO-US023328.
PR 08-NOV-2000; 2000WO-US030952.
PR 10-NOV-2000; 2000WO-US030873.
PR 01-DEC-2000; 2000WO-US032678.
PR 20-DEC-2000; 2000US-00747259.
PR 28-FEB-2001; 2000WO-US034956.
PR 28-FEB-2001; 2001US-00796498.

PR 28-FEB-2001; 2001WO-US006520.
PR 01-MAR-2001; 2001WO-US006666.
PR 09-MAR-2001; 2001US-00802706.
PR 14-MAR-2001; 2001US-00808689.
PR 22-MAR-2001; 2001US-00816744.
PR 05-APR-2001; 2001US-00828366.
PR 10-MAY-2001; 2001US-00854208.
PR 18-MAY-2001; 2001US-00860216.
PR 25-MAY-2001; 2001US-00866028.
PR 25-MAY-2001; 2001US-00866034.
PR 25-MAY-2001; 2001WO-US017092.
PR 01-JUN-2001; 2001US-00872035.
PR 01-JUN-2001; 2001WO-US017800.
PR 05-JUN-2001; 2001US-00874503.
PR 14-JUN-2001; 2001US-00882636.
PR 19-JUN-2001; 2001US-00886342.
PR 20-JUN-2001; 2001WO-US019692.
PR 21-JUN-2001; 2001US-00887879.
PR 22-JUN-2001; 2001WO-US020115.
PR 29-JUN-2001; 2001WO-US021066.
PR 09-JUL-2001; 2001WO-US021735.
PR 18-JUL-2001; 2001US-00908827.
PR 06-AUG-2001; 2001US-00924419.
PR 09-AUG-2001; 2001US-00927796.
PR 16-AUG-2001; 2001US-00931836.
PR 19-DEC-2001; 2001US-00028072.

(GETH) GENENTECH INC.

Baker KP, Beresini M, Deforge L, Desnoyers L, Filvaroff E, Gao W;
Gerritsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S;
Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;

WPI; 2003-687639/65.
N-PSDB; ADA76121.

New isolated nucleic acid encoding a secreted and transmembrane polypeptide, designated e.g. PRO1114 or PRO4978, useful in chromosome and gene mapping, in generating antisense RNA and DNA, and in gene therapy.

Claim 12; Fig 172; 659pp; English.

The invention relates to isolated human PRO polypeptides (secreted and transmembrane polypeptides) and the polynucleotides encoding them. The invention also relates to an antibody which specifically binds to a PRO polypeptide, a method for stimulating the release of tumour necrosis factor-alpha (TNF-alpha) from human blood, a method for stimulating the proliferation or differentiation of chondrocyte cells and a method for detecting the presence of a tumour in a mammal (e.g. adrenal, lung, colon, breast, prostate, rectal, kidney, cervical and liver tumours). The polynucleotides are useful in molecular biology, including uses as hybridisation probes, in chromosome and gene mapping, in generating antisense RNA and DNA and in gene therapy. The polynucleotides may also be used in preparing PRO polypeptides by recombinant techniques and in generating either transgenic animals or knock-out animals which are useful in the development and screening of therapeutically useful reagents. The PRO polypeptides or antibodies are used in preparing a medicament for treating a condition responsive to the polypeptides or antibodies, such as tumours, for stimulating and inhibiting proliferation of human microvascular endothelial cells, for modulating the uptake of glucose or FFA by skeletal muscle cells or adipocyte cells, for stimulating differentiation of adipocyte cells, for stimulating proliferation of or gene expression in pericyte cells, for stimulating the proliferation of inner ear utricular supporting cells or T-lymphocyte cells, for inducing endothelial cell tube formation and for treating various bone and/or cartilage disorders such as sports injuries and arthritis. PRO polypeptides which stimulate the release of proteoglycans from cartilage are useful for treating sports-related joint problems, articular cartilage defects, osteoarthritis and rheumatoid arthritis. PRO polypeptides are also useful for treating various mammalian haemoglobin-associated disorders such as various thalassaemias and conditions which may benefit from enhanced local immune system cell infiltration. This

CC sequence represents a human PRO polypeptide of the invention. Note: The
CC sequence data for this patent is also available in electronic format from
CC USPTO at seqdata.uspto.gov/sequence.html.
XX
SQ Sequence 182 AA;

Query Match: 99.7%; Score 959; DB 6; Length 182;
Best Local Similarity 99.5%; Pred. No. 7.6e-95;
Matches 181; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MEPQLGPEAAALRPGWLALLLWVSALSCSFSLPASSLSLVPQVRTSYNFGRTFLGLDKC 60
Db 1 MEPQLGPEAAALRPGWLALLLWVSALSCSFSLPASSLSLVPQVRTSYNFGRTFLGLDKC 60

Qy 61 NACIGTSICKKFFKEIRSDNWLASHLGLPPDSLLSYPNYSDDSKIMRPVEIFRLVSKY 120
Db 61 NACIGTSICKKFFKEIRSDNWLASHLGLPPDSLLSYPNYSDDSKIMRPVEIFRLVSKY 120

Qy 121 QNEISDRKICASASAPKTCSTIERVLRKTERFQKWLQAKRLTPDLVQDCHQGORELKFLCM 180
Db 121 QNEISDRRIRICASASAPKTCSTIERVLRKTERFQKWLQAKRLTPDLVQDCHQGORELKFLCM 180

Qy 181 LR 182
Db 181 LR 182

Search completed: June 8, 2004, 16:24:24
Job time : 60 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 8, 2004, 16:22:14 ; Search time 20 Seconds
(without alignments)
875.343 Million cell updates/sec

Title: US-10-054-988-114
Perfect score: 962
Sequence: 1 MEPQLSPBAALRPGWLALL.....DLVQDCHQSQRELKFLCMLR 182

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues
Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78:*
1: Pirl:*
2: Pirl2:*
3: Pirl3:*
4: Pirl4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	86.5	9.0	438	T32114	hypothetical prote
2	83	8.6	398	F91202	EspG protein [impo
3	83	8.6	398	A86049	hypothetical prote
4	81	8.4	424	AB1034	UV protection prot
5	80	8.3	390	S56560	hypothetical 43.6K
6	80	8.3	390	B91291	hypothetical prote
7	80	8.3	390	D86132	hypothetical prote
8	79	8.2	424	B38176	samB protein - Sal
9	79	8.2	985	T41135	hypothetical prote
10	77	8.0	678	T50256	probable vacuolar
11	77	8.0	2374	T21052	hypothetical prote
12	76	7.9	820	T26272	hypothetical prote
13	76	7.9	1006	E96683	hypothetical prote
14	76	7.9	2712	T30949	hypothetical prote
15	76	7.9	3864	D87757	hypothetical prote
16	75	7.8	252	C70380	protein C44E4.1a [
17	75	7.8	386	G88765	hypothetical prote
18	75	7.8	427	C86469	protein T11B7.1 [i
19	75	7.8	445	T24829	hypothetical prote
20	74.5	7.7	614	T39688	fibrin - fission
21	74.5	7.7	795	T05860	alpha,alpha-trehal
22	74.5	7.7	1963	B98002	IgA-specific metal
23	74.5	7.7	2004	F95133	immunoglobulin A1
24	74	7.7	398	T26274	hypothetical prote
25	74	7.7	464	F97714	hypothetical prote
26	74	7.7	481	T00863	hypothetical prote
27	74	7.7	600	T21231	hypothetical prote
28	74	7.7	739	T25030	hypothetical prote
29	74	7.7	1109	T18536	receptor-like prot

30	74	7.7	1144	2	D97227	pyruvate carboxyla
31	74	7.7	1175	2	T46124	hypothetical prote
32	73.5	7.6	342	2	A56552	homeotic protein H
33	73.5	7.6	1447	2	S50918	DNA helicase TPS1
34	73.5	7.6	1839	1	RRWPEM	genome polyprotein
35	73	7.6	1454	2	B84535	probable retroelem
36	73	7.6	1465	2	S43529	165K protein, skel
37	72.5	7.5	167	2	S46656	ubiquitin-protein
38	72.5	7.5	840	2	G98169	hypothetical prote
39	72.5	7.5	840	2	AF3117	hypothetical prote
40	72	7.5	424	2	JQ0661	impB protein - Sal
41	72	7.5	461	2	H71717	3-deoxy-d-manno-oc
42	72	7.5	545	2	T01288	protein kinase F27
43	72	7.5	588	2	T31675	hypothetical prote
44	72	7.5	1020	2	G88208	protein K02A2.3 [i
45	72	7.5	1071	2	T18307	suppressor protein

ALIGNMENTS

RESULT 1

T32114 hypothetical protein F59E11.11 - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999

C;Accession: T32114

R;Bradshaw, H.

submitted to the EMBL Data Library, July 1997

A;Description: The sequence of C. elegans cosmid F59E11.

A;Reference number: Z21124

A;Accession: T32114

A;Status: preliminary; translated from GB/EMBL/DBBJ

A;Molecule type: DNA

A;Residues: 1-438 <BRA>

A;Cross-references: EMBL:AF016685; PIDN:AB66223.1; GSPDB:GN00023; CESP:F59E11.11

A;Experimental source: strain Bristol N2; clone F59E11

C;Genetics:

A;Gene: CESP:F59E11.11

A;Map position: 5

A;Introns: 51/2; 75/2; 141/3; 184/3; 214/3; 285/3; 317/2; 356/1

Query Match 9.0%; Score 86.5; DB 2; Length 438;
Best Local Similarity 24.1%; Pred. No. 2.4;
Matches 42; Conservative 26; Mismatches 49; Indels 57; Gaps 11;

QY 26 LSCSFSLPASSLSLVQVRTSYNF-----GRTFLGLDKCNAC-----IGTS 67

Db 1 MSCATS-PVSSSESSIGSVESNFDLKEICCSKGHNHF-GVTACACAAFFRRMIVGTG 58

QY 68 I---CKKPFKEIRSDNWLAS-----HLGLPPDS-----LLSYPNYSD---DSKI 107

Db 59 YRQKCRSYNTCEPKDGRWQCKKRLKCYSLGTMTPDSKSQLTMRILYLYSDIQHDDL 118

QY 108 WRPVEIFRLVSKYQNEISDRKICASASAPKTCISIERVL-----RKTERFQK 153

Db 119 FQCSLIFREKNRKRKLIE-----NIPKT--VEKILGIPHLILFRKKERLHE 163

RESULT 2

F91202

EspG protein [imported] - Escherichia coli (strain O157:H7, substrain RIMD 0509952)

C;Species: Escherichia coli

C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 18-Jul-2001

C;Accession: F91202

R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.

Gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.

DNA Res. 8, 11-22, 2001

A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and ge

A;Reference number: A99629; MUID:21156231; PMID:11258796

A;Accession: F91202

A;Status: preliminary

A;Molecule type: DNA


```
Db 188 ANFYHLGQLNPALSGSDILKVYGVATFRDKEALINELDAMTARVQQWEEGQRLDPRP 247
QY 164 --LVQDCHQG 171
Db 248 RLITGCPIG 257

RESULT 6
B91291
hypothetical protein ECS5298 [imported] - Escherichia coli (strain O157:H7, substrain R)
C:Species: Escherichia coli
C>Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Aug-2001
C:Accession: B91291
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
Gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and geno
A:Reference number: A99629; MUID:21156231; PMID:11258796
A:Accession: B91291
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-390 <EAY>
A:Cross-references: GB:BA000007; PIDN:BA038721.1; PID:gl3364776; GSPDB:GN00154
A:Experimental source: strain O157:H7, substrain RMD 0509952
C:Genetics:
A:Gene: ECS5298
C:Superfamily: 2-hydroxyglutaryl-CoA dehydratase beta chain

Query Match 8.3%; Score 80; DB 2; Length 390;
Best Local Similarity 22.1%; Pred. No. 8.6;
Matches 42; Conservative 31; Mismatches 45; Indels 72; Gaps 13;

QY 39 SLVPQVRTSYNFGRTFLGLDKC-----NACIGTSIC---KKFKEEIRSDNWLASHLGL 89
Db 83 NLCPLIKSSYGFGKT---DKCPYFYFSDLVVGETTCDGKKMYE-----YMAE---F 128

QY 90 PPDSLLSYYPANYSDDSK--IWRPVEIFRL----VSKYQNEIS-----DRKIC 130
Db 129 KPVHVMQLPNSVKDDASRALWK-AEMLRLQKAVEERFGHEISEDALRDAIALKNRERRAL 187

QY 131 AS-----ASAPKTCIERVLR---KTERF-----QKWLOAKRLTPD- 163
Db 188 ANFYHLGQFNPPALSGSDILKVYGVATFRDKEALINELDAMTARVQQWEEGQRLDPRP 247

QY 164 --LVQDCHQG 171
Db 248 RLITGCPIG 257

RESULT 7
D86132
hypothetical protein yjiM [imported] - Escherichia coli (strain O157:H7, substrain EDL93
C:Species: Escherichia coli
C>Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
C:Accession: D86132
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: D86132
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-390 <STO>
A:Cross-references: GB:AE005174; NID:gl2519354; PIDN:AAGS9520.1; GSPDB:GN00145; UWGP:Z59
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: yjiM
C:Superfamily: 2-hydroxyglutaryl-CoA dehydratase beta chain

Query Match 8.3%; Score 80; DB 2; Length 390;
Best Local Similarity 22.1%; Pred. No. 8.6;
```

```
Matches 42; Conservative 31; Mismatches 45; Indels 72; Gaps 13;

QY 39 SLVPQVRTSYNFGRTFLGLDKC-----NACIGTSIC---KKFKEEIRSDNWLASHLGL 89
Db 83 NLCPLIKSSYGFGKT---DKCPYFYFSDLVVGETTCDGKKMYE-----YMAE---F 128

QY 90 PPDSLLSYYPANYSDDSK--IWRPVEIFRL----VSKYQNEIS-----DRKIC 130
Db 129 KPVHVMQLPNSVKDDASRALWK-AEMLRLQKAVEERFGHEISEDALRDAIALKNRERRAL 187

QY 131 AS-----ASAPKTCIERVLR---KTERF-----QKWLOAKRLTPD- 163
Db 188 ANFYHLGQFNPPALSGSDILKVYGVATFRDKEALINELDAMTARVQQWEEGQRLDPRP 247

QY 164 --LVQDCHQG 171
Db 248 RLITGCPIG 257

RESULT 8
B38176
samB protein - Salmonella typhimurium
C:Species: Salmonella typhimurium
C>Date: 28-Aug-1992 #sequence_revision 28-Aug-1992 #text_change 20-Jun-2000
C:Accession: B38176
R;Nohmi, T.; Nakai, A.; Nakai, Y.; Watanabe, M.; Murayama, S.Y.; Sofuni, T.
J. Bacteriol. 173, 1051-1063, 1991
A:Title: Salmonella typhimurium has two homologous but different umuDC operons: cloni
A:Reference number: A38176; MUID:91123176; PMID:1991707
A:Accession: B38176
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-424 <NOH>
A:Cross-references: GB:D90202; NID:g217087; PIDN:BAAL4226.1; PID:g217089
A:Experimental source: strain LT2
C:Genetics:
A:Gene: samB
C:Function:
A:Description: restores UV mutability; involved in mutagenesis
C:Superfamily: umuC protein
C:Keywords: DNA repair; induced mutagenesis; SOS mutagenesis

Query Match 8.2%; Score 79; DB 2; Length 424;
Best Local Similarity 24.2%; Pred. No. 12;
Matches 40; Conservative 22; Mismatches 63; Indels 40; Gaps 6;

QY 23 VSALSCSFSLPAS-----SLSLVPQVRTSYNFGRTFLGLDKCNACIGTSICKKFFKE 75
Db 67 VIAFSSNYALYASMSNRVMVHLEELAPVE-QYSIDEMFLDIRGIDSCIDFEDFGRQLRE 125

QY 76 EIRSDNWLASHLGLPPDPSLLSYYPANYSDDSKIM-----RPVE 112
Db 126 HVRSGTGLTIGVGMGPTKTLAKSAQWA--SKWSQFGVGLALTLHNQKTEKLLSLQPVE 183

QY 113 ----IFRLVSKYQNEISDRKICASASAPKTC---SIERVLRKTER 150
Db 184 EIWVGRRISKXLTMTGITTALQLARANPTFIRKNFNVVLTVR 228

RESULT 9
T41135
hypothetical protein SPCC1795.08c - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C:Accession: T41135
R;Lyne, M.; Rajandream, M.A.; Barrell, B.G.; Oliver, K.; Harris, D.
submitted to the EMBL Data Library, April 1998
A:Reference number: Z21971
A:Accession: T41135
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-985 <LYN>
A:Cross-references: EMBL:AL022598; PIDN:CAA18643.1; GSPDB:GN00068; SPDB:SPCC1795.08c
```


QY 139 CSIERVLRKTERFQXW 154
Db 808 ASVE---KGMDDFADW 820

RESULT 13
E96683
hypothetical protein F12P19.9 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C;Accession: E96683
R;Theologos, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: E96683
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1006 <STO>
A;Cross-references: GB:AE005173; NID:G6227017; PIDN:AAF06053.1; GSPDB:GN00141
C;Genetics:
A;Gene: F12P19.9
A;Map position: 1

Query Match 7.9%; Score 76; DB 2; Length 1006;
Best Local Similarity 26.7%; Pred. No. 64;
Matches 31; Conservative 17; Mismatches 48; Indels 20; Gaps 5;

QY 24 SALSCSFSLPASS--LSSLVPQVTSYNGRTFLGLKGNACIGTSICKKFFKEIRSDN 81
Db 190 NGLSCSSDSFAESDALGPVSSYETDYDFRNS-----DCRSTGSELCR-----FSSQR 238

QY 82 WLASHLGLPPDLSLLSP---ANYSDDSKIMRPVEIFRLVSKYQNEISDRKICASAS 134
Db 239 FAAS-----PPLSIITQPVTRSNVLKDIMINGAITGLIDGSKNQNDALSPKLLSAT 290

RESULT 14
T30949
hypothetical protein C44E4.1a - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 22-Oct-1999
C;Accession: T30949
R;Sammons, L.; Wohldmann, P.; Gillam, B.
submitted to the EMBL Data Library, August 1999
A;Description: The sequence of C. elegans cosmid C44E4.
A;Reference number: Z20945
A;Accession: T30949
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-2712 <SAM>
A;Cross-references: EMBL:AF003140; PIDN:AAD47119.1
A;Experimental source: strain Bristol N2; clone C44E4
C;Genetics:
A;Map position: 1
A;Introns: 44/1; 346/2; 870/1; 1005/3; 1404/3; 1506/2; 1665/1; 1891/3; 1936/3; 2068/2; 2
A;Note: C44E4.1a

Query Match 7.9%; Score 76; DB 2; Length 2712;
Best Local Similarity 25.0%; Pred. No. 2.1e+02;
Matches 27; Conservative 19; Mismatches 34; Indels 28; Gaps 5;

QY 83 LASHLGLPPD-SLLSYPPANYSDSKIMRPVEIFRLVSKYQNEISDR-----KIC 130
Db 856 LVSSSANPNTSIMNWPST-SEDSYIIACTDLILLIPQHLQELDRRKSVPRDDQWIKLC 914

QY 131 ASASAPKTCIERVLRKTERFQKWLQAKRLTPDLVQDCHQOQRELKFL 178
Db 915 QLASLSSGCSAYR-----QCKKL---LLAMCHNDENKYKIM 947

RESULT 15
D87757
protein C44E4.1a [imported] - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001
C;Accession: D87757
R;anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A;Title: Genome sequence of the nematode C. elegans: a platform for investigating bio-
A;Reference number: A75000; MUID:99069613; PMID:9851916
A;Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/projects/C_e-
A;Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999;
A;Accession: D87757
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-3864 <STO>
A;Cross-references: GB:chr_I; PIDN:AAB54167.1; PID:G2088725; GSPDB:GN00019; CESP:C44E4
C;Genetics:
A;Gene: C44E4.1a
A;Map position: 1

Query Match 7.9%; Score 76; DB 2; Length 3864;
Best Local Similarity 25.0%; Pred. No. 3.3e+02;
Matches 27; Conservative 19; Mismatches 34; Indels 28; Gaps 5;

QY 83 LASHLGLPPD-SLLSYPPANYSDSKIMRPVEIFRLVSKYQNEISDR-----KIC 130
Db 1993 LVSSSANPNTSIMNWPST-SEDSYIIACTDLILLIPQHLQELDRRKSVPRDDQWIKLC 2051

QY 131 ASASAPKTCIERVLRKTERFQKWLQAKRLTPDLVQDCHQOQRELKFL 178
Db 2052 QLASLSSGCSAYR-----QCKKL---LLAMCHNDENKYKIM 2084

Search completed: June 8, 2004, 16:26:20
Job time : 21 secs

QY 126 DRK-----ICASASAPKTCSTI--ERVLRKTERFQKWLQ 156
Db 183 QQKGDSPYQILQFAIAPKORSILHDRI---ALRFQKMIIE 219

RESULT 2
YJIM_ECOLI
ID YJIM_ECOLI STANDARD; PRT; 383 AA.
AC P39384;
DT 01-FEB-1995 (Rel. 31, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein yjiM.
GN YJIM OR B4335.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=KL2 / MG1655;
RX MEDLINE=95334362; PubMed=7610040;
RA Burland V.D., Plunkett G. III, Sofia H.J., Daniels D.L.,
RA Blattner F.R.;
RT "Analysis of the Escherichia coli genome VI: DNA sequence of the
RT region from 92.8 through 100 minutes.";
RL Nucleic Acids Res. 23:2105-2119(1995).
CC -!- SIMILARITY: STRONG, TO M.JANNASCHII MJ0007.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; U14003; AAA97231.1; ALT INIT.
DR EMBL; AE000504; AAC77291.1; ALT_INIT.
DR EcoGene; EGI2574; yjiM.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 383 AA; 42742 MW; 701CB569D0BFDACC CRC64;

Query Match 8.3%; Score 80; DB 1; Length 383;
Best Local Similarity 22.1%; Pred. No. 2.9;
Matches 42; Conservative 31; Mismatches 45; Indels 72; Gaps 13;

QY 39 SLVPQVTSYNGRTFLGLDKC-----NACIGTSIC--KKFFKEIRSDNWLASHLGL 89
Db 76 NLCPKLSYGFQKT---DKCPYFYFSDLVVGETTCDGKKQMYE-----YMAE--F 121
QY 90 PPDLSLSYPANYSDDSK--IWRPVEIFRJ-----VSKYQNEIS-----DRKIC 130
Db 122 KPVHVMQLPNSVKDDASRALWK-AEMLRLQKTVEERFGHEISEDALRDAIAKNRERRAL 180
QY 131 AS-----ASAPKTCSTIERVLR-----KTERF-----QKWLQAKRLTPD- 163
Db 181 ANFYHGLQNPALSGSDILKVVYGATERFDKEALINELDAMTARVRQWEEGQRLDPRP 240

QY 164 --LVQDCCHQ 171
Db 241 RLITGCPIG 250

RESULT 3
SAMB_SALTY
ID SAMB_SALTY STANDARD; PRT; 424 AA.
AC P23832;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Samb protein.
GN SAMB.

OS Salmonella typhimurium.
OG Plasmid 60-mDa cryptic.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=602;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LT2;
RX MEDLINE=91123176; PubMed=1991707;
RA Nohmi T., Hakura A., Nakai Y., Watanabe M., Murayama S.Y.,
RA Sofuni T.;
RT "Salmonella typhimurium has two homologous but different umuDC
RT operons: cloning of a new umuDC-like operon (samAB) present in a
RT 60-megadalton cryptic plasmid of S. typhimurium.";
RL J. Bacteriol. 173:1051-1063(1991).
CC -!- FUNCTION: Involved in UV protection and mutation.
CC -!- SIMILARITY: Belongs to the DNA polymerase type-Y family.
CC -!- SIMILARITY: Contains 1 umuC domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; D90202; BAA14226.1; -.
DR PIR; B38176; B38176.
DR HAMAP; MF_01113; atypical; 1.
DR InterPro; IPR001126; UMuC_like.
DR Pfam; PF00817; IMS; 1.
DR PROSITE; PS50173; UMuC; 1.
KW Plasmid; SOS mutagenesis; DNA repair.
FT DOMAIN 2 189 UMuC.
SQ SEQUENCE 424 AA; 47727 MW; PF8C47476CC58A2B CRC64;

Query Match 8.2%; Score 79; DB 1; Length 424;
Best Local Similarity 24.2%; Pred. No. 4.2;
Matches 40; Conservative 22; Mismatches 63; Indels 40; Gaps 6;

QY 23 VSALSCSFSLPAS-----SLSLVQVTSYNGRTFLGLDKCNACIGTSICKKFFKE 75
Db 67 VIAFSSNYALYASMSNRVMHLEAPRVE-QYSIDEMFLDIRGIDSCIDFDFGRQLRE 125
QY 76 EIRSDNWLASHLGLPPDLSLSYPANYSDDSKIW-----RPVE 112
Db 126 HVRSGTGLTIGVGMGPTKTLAKSAQWA--SKEWSQGGVLAALTLHNQKTEKLLSLQPVE 183
QY 113 ----IFRLVSKYQNEISDRKICASASAPKTC---SIERVLRKTER 150
Db 184 EIMGVGRIRISKINTMGITTALQLARANPTFIRKNFNVVLTVR 228

RESULT 4
VPS1_SCHPO
ID VPS1_SCHPO STANDARD; PRT; 678 AA.
AC Q9URZ5; O14309;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Vacuolar sorting protein 1.
GN VPS1 OR SPAC767.01C OR SPAC9G1.14C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,


```
FT DOMAIN 97 117 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 118 139 3 (POTENTIAL).
FT DOMAIN 140 162 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 163 183 4 (POTENTIAL).
FT DOMAIN 184 211 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 212 235 5 (POTENTIAL).
FT DOMAIN 236 263 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 264 285 6 (POTENTIAL).
FT DOMAIN 286 302 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 303 326 7 (POTENTIAL).
FT DOMAIN 327 366 CYTOPLASMIC (POTENTIAL).
FT DISULFID 116 198 BY SIMILARITY.
FT CARBOHYD 13 13 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 27 27 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARSPLIC 266 289 AVVFAFILCWLPFHVGRYLFSSK -> GGSQCALELSLPG
PLHSSCLFSSP (in isoform 1B).
/FTID=VSP 001918.
Missing (in isoform 1B).
/FTID=VSP 001919.
FT VARSPLIC 290 366
SQ SEQUENCE 366 AA; 41194 MW; 2C850B3EF61B7C1C CRC64;
```

Query Match 7.8%; Score 75; DB 1; Length 366;
Best Local Similarity 25.0%; Pred. No. 8.5;
Matches 39; Conservative 23; Mismatches 46; Indels 48; Gaps 8;

```
QY 11 ALRPGWLALLMVSALSCFSLPASSLSLVQVRTSYNFGRTFLGLDKCNACIGTSICK 70
Db 204 AVRSGLLTVMVWVS--SVFFLPVFLTLVLSLI-----GRKLWRRKRGSAVGSRLRD 255
QY 71 KFFKEEIRSDN-----WLASHLG-----LPPDS-----LLSY----- 97
Db 256 QNHKQTVKMLAVVVFAPILCWLPFHVGRYLFSSKLEPGSVEIAQISQYCNLVSFVLYLS 315
QY 98 ----PANYSDDSKIWRPVEIFRLVSKYQNEISDRKI 129
Db 316 AAINPILYNIMSKYR-VAVEKLIG--FEFESQKRL 348
```

```
RESULT 9
Y140 HUMAN
ID Y140 HUMAN STANDARD; PRT; 422 AA.
AC Q14153;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein KIAA0140.
GN KIAA0140.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Bone marrow;
RX MEDLINE=96127530; PubMed=8590280;
RA Nagase T., Seki N., Tanaka A., Ishikawa K.-I., Nomura N.;
RT "Prediction of the coding sequences of unidentified human genes. IV.
RT The coding sequences of 40 new genes (KIAA0121-KIAA0160) deduced by
RT analysis of cDNA clones from human cell line KG-1.";
RL DNA Res. 2:167-174(1995).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; D50930; BAA09489.1; -
CC Hypothetical protein.
KW
SQ SEQUENCE 422 AA; 45796 MW; 8F607F093408CEAC CRC64;
```

```
Query Match 7.7%; Score 74.5; DB 1; Length 422;
Best Local Similarity 23.3%; Pred. No. 11;
Matches 30; Conservative 19; Mismatches 45; Indels 35; Gaps 5;

QY 33 PASSLSLVQVRTSYNFGRTFLGLDKCNACIGTSICKKPFKEEIRSDNWLASHLG---L 89
Db 63 PSTSIWECLPEKDS-----LWHREAVTACAVTSLIK-----DLSISDHNGNPSA 107
QY 90 PPDLSLSYPANYSDD-----SKWRPVEIFRL-----VSKYQNEISDRKICAS 132
Db 108 PPSKRQCRSLSPFSDMSRCTSRPLGSKVWTPVEKRCYSGSVQRYSGFSTMQRSSS 167
QY 133 ASAPKTCISI 141
Db 168 FSLPSRANV 176
```

RESULT 10

```
OV01 MOUSE
ID OV01 MOUSE STANDARD; PRT; 267 AA.
AC Q9WTJ2;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Putative transcription factor Ovo-like 1 (mOv01) (mOv01a).
GN OV01.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Skin;
RA Schonbaum C.P., Fantes J., Mahowald A.P.;
RT "Characterization of mouse and Caenorhabditis elegans genes related to
RT the Drosophila melanogaster ovo/svb gene.";
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP FUNCTION.
RX MEDLINE=99026118; PubMed=9808631;
RA Dai X., Schonbaum C., Degenstein L., Bai W., Mahowald A., Fuchs E.;
RT "The ovo gene required for cuticle formation and oogenesis in flies is
RT involved in hair formation and spermatogenesis in mice.";
RL Genes Dev. 12:3452-3463(1998).
CC -!- FUNCTION: Putative transcription factor. Involved in hair
CC formation and spermatogenesis. May function in the differentiation
CC and/or maintenance of the urogenital system.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- TISSUE SPECIFICITY: EXPRESSED IN SKIN, TESTIS, KIDNEY AND WEAKLY
CC IN LUNG. NOT DETECTED IN HEART, BRAIN, SPLEEN, LIVER AND SKELETAL
CC MUSCLE.
CC -!- DEVELOPMENTAL STAGE: FIRST EXPRESSED AT E14.5 DAY IN THE
CC SUPRABASAL LAYERS OF DEVELOPING EPIDERMIS, AT E15.5 EXPRESSION
CC BEGINS IN THE INNER CELLS OF DEVELOPING HAIR GERMS AND RESTRICTED
CC TO INNER ROOT SHEATH AND/OR PRECORTICAL CELLS OF DEVELOPING HAIR
CC FOLLICLES.
CC -!- SIMILARITY: Contains 4 C2H2-type zinc fingers.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF134804; AAD29689.1; -
CC EMBL; AF134805; AAD29690.1; -
CC HSSP; P07248; 2ADR.
CC MGD; MGI:1330290; Ov01.
CC GO; GO:0005634; C:nucleus; IDA.
CC GO; GO:0008544; P:epidermal differentiation; IMP.
CC GO; GO:0007498; P:mesoderm development; IMP.
DR
```



```
DR GO:0007283; P: spermatogenesis; IMP.
DR InterPro; IPR007087; Znf_C2H2.
DR Pfam; PF00096; zf-C2H2; 4.
DR ProDom; PDC00003; Znf_C2H2; 1.
DR SMART; SM00355; Znf_C2H2; 4.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 3.
DR PROSITE; PS50157; ZINC_FINGER_C2H2_2; 3.
KW Zinc-finger; Metal-binding; DNA-binding; Nuclear protein; Repeat;
KW Transcription regulation.
FT ZN_FING 118 140 C2H2-TYPE 1.
FT ZN_FING 146 168 C2H2-TYPE 2.
FT ZN_FING 174 197 C2H2-TYPE 3.
FT ZN_FING 213 236 C2H2-TYPE 4.
SQ SEQUENCE 267 AA; 30221 MW; DAD4F51150C21C2D CRC64;

Query Match
Best Local Similarity 7.7%; Score 74; DB 1; Length 267;
Matches 37; Conservative 23; Mismatches 81; Indels 30; Gaps 4;

QY 2 EPQLGPERAALRPGWLALLLWVSALSCFS-----LPSASSLSLVPQVRTSYNFG 51
Db 42 QPYREPAASVAPPPSCPLALDMSLRDSSYSVAPGCVVAQLPSEDVSHLTDPQSRDQGF 101

QY 52 RTFLGLDKCNACIGTS-----ICKKFFKEIRSDNWLASHLGLPPDLSLLSYPNYS 102
Db 102 RT-----KMKVTLGDSPNGDLFTCHICQKSFTHQMLNLRHMKCHNDVKRHLCTYCGKGF 156

QY 103 DSKIWPRVEIFRLVSKYQNEISDRKICASASAPKTCISIERVLRKTERFQK 153
Db 157 DTFDLKRVHRTHTGVRPYKCSLCKD-----AFTQRCLESLEHLKTHGVQQ 201

RESULT 11
HKCA_MOUSE
ID HKCA_MOUSE STANDARD; PRT; 342 AA.
AC P31257;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Homeobox protein Hox-C10 (Hox-3.6).
GN HOXC10 OR HOXC-10 OR HOX-3.6.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CD-1;
RX MEDLINE=92360500; PubMed=1353983;
RA Peterson R.L., Jacobs D.F., Awgulewitsch A.;
RT "Hox-3.6: isolation and characterization of a new murine homeobox
RT gene located in the 5' region of the Hox-3 cluster.";
RL Mech. Dev. 37:151-166 (1992).
RN [2]
RP SEQUENCE OF 289-313 FROM N.A.
RX MEDLINE=92073357; PubMed=1720547;
RA Murtha M.T., Leckman J.F., Ruddle F.H.;
RT "Detection of homeobox genes in development and evolution.";
RL Proc. Natl. Acad. Sci. U.S.A. 88:10711-10715 (1991).
CC -!- FUNCTION: Sequence-specific transcription factor which is part of
CC a developmental regulatory system that provides cells with
CC specific positional identities on the anterior-posterior axis.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- SIMILARITY: Belongs to the Abd-B homeobox family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

```
DR EMBL; X63507; -; NOT ANNOTATED_CDS.
DR EMBL; M81658; AAA63310.1; -.
DR PIR; A56552; A56552.
DR HSSP; P02833; 9ANT.
DR TRANSPAC; T03350; -.
DR MGD; MGI:96192; Hoxc10.
DR InterPro; IPR001356; Homeobox.
DR Pfam; PF00046; homeobox; 1.
DR PRINTS; PR00024; HOMEBOX.
DR ProDom; PD000010; Homeobox; 1.
DR SMART; SM00389; HOX; 1.
DR PROSITE; PS00027; HOMEBOX 1; 1.
DR PROSITE; PS50071; HOMEBOX_2; 1.
KW Homeobox; DNA-binding; Developmental protein; Nuclear protein;
KW Transcription regulation.
FT DNA_BIND 268 327 HOMEBOX.
SQ SEQUENCE 342 AA; 38195 MW; 7BB2E117A768F52B CRC64;

Query Match
Best Local Similarity 7.6%; Score 73.5; DB 1; Length 342;
Matches 42; Conservative 15; Mismatches 74; Indels 35; Gaps 6;

QY 6 GPERAALRPGWLALLLWVSALSCFSFSLPSSLSLVPQVRTSYNFGRTFLGLDKCNACIG 65
Db 123 GPEALY-----SHPLPESCLGEHEVPVPSYRASPSYSALDKTPHCAG 166

QY 66 TSICKKFFKEIRSDNWLASHLGLPP-DSLISYPANYSDDSKIWPRVEIFRLVSKYQNEI 124
Db 167 ANEFEAPF-EQASLNPRTEHLESPLQGGKVSFFETPKSDSQTPSPNEI-----KTEQSL 220

QY 125 SDRKICASA-----SAPKTCISIE--RVLKTERFQKWLQAK 158
Db 221 AGPKASPSSESEKERAKTADSSPTSDNEAKEIKAEKNTGNWLTAK 266

RESULT 12
SGS1_YEAST
ID SGS1_YEAST STANDARD; PRT; 1447 AA.
AC P35187;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Helicase SGS1 (Helicase Tps1).
GN SGS1 OR TPS1 OR YMR190C OR YMR646.02C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=W303;
RX MEDLINE=95059068; PubMed=7969174;
RA Gangloff S., McDonald J.P., Bendixen C., Arthur L., Rothstein R.;
RT "The yeast type I topoisomerase Top3 interacts with Sgs1, a DNA
RT helicase homolog: a potential eukaryotic reverse gyrase.";
RL Mol. Cell. Biol. 14:8391-8398 (1994).
RN [2]
RP SEQUENCE FROM N.A.
RA Romeo A.M., Kleff S., Sternglanz R.;
RL Submitted (DEC-1992) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RX STRAIN=S288C / AB972;
RX MEDLINE=97313268; PubMed=9169872;
RA Bowman S., Churcher C.M., Badcock K., Brown D., Chillingworth T.,
RA Connor R., Dedman K., Devlin K., Gentles S., Hamlin N., Hunt S.,
RA Jagels K., Iye G., Moule S., Odell C., Pearson D., Rajandream M.A.,
RA Rice P., Skelton J., Walsh S., Whitehead S., Barrell B.G.;
RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome
RT XIII.";
RL Nature 387:90-93 (1997).
RN [4]
RP STRUCTURE BY NMR OF HRDC DOMAIN.
```

```
RX MEDLINE=20113483; PubMed=10647186;
RA Liu Z., Macias M.J., Bottomley M.J., Stier G., Linge J.P., Nilges M.,
RA Bork P., Sattler M.;
RT "The three-dimensional structure of the HRDC domain and implications
RL for the Werner and Bloom syndrome proteins.";
RL Structure 7:1557-1566(1999).
CC -!- FUNCTION: Interacts with topoisomerases II and TOP3. Could create
CC a deleterious topological substrate that TOP3 preferentially
CC resolves. The TOP3-SGS1 protein complex may function as a
CC eukaryotic reverse gyrase introducing positive supercoils into
CC extrachromosomal ribosomal DNA rings.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- SIMILARITY: Belongs to the helicase family. RecQ subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement [See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch].
CC -----
DR EMBL; L07870; AAA35167.1; -
DR EMBL; U22341; AAB60289.1; -
DR EMBL; Z47815; CAAB7811.1; -
DR PIR; S50918; S50918.
DR PDB; 1D8B; 10-JAN-00.
DR GerMOnline; 142864; -
DR SGD; S0004802; SGS1.
DR GO; GO:0005730; C:nucleolus; IDA.
DR GO; GO:0004003; F:ATP dependent DNA helicase activity; IDA.
DR GO; GO:0007001; P:chromosome organization and biogenesis [en. . .; IMP.
DR GO; GO:0006268; P:DNA unwinding; IDA.
DR GO; GO:0045132; P:meiotic chromosome segregation; IMP.
DR GO; GO:0000070; P:mitotic chromosome segregation; IMP.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR002464; DEAD_box.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR002121; HRDC.
DR InterPro; IPR004589; RecQ.
DR Pfam; PF00270; DEAD; 1.
DR Pfam; PF00271; helicase_C; 1.
DR Pfam; PF00570; HRDC; 1.
DR SMART; SM00487; DEXdc; 1.
DR SMART; SM00490; HELIC; 1.
DR SMART; SM00341; HRDC; 1.
DR TIGRfams; TIGR00614; recQ_fam; 1.
DR PROSITE; PS00690; DEAD_ATP_HELICASE; UNKNOWN_1.
KW Helicase; ATP-binding; Nuclear protein; 3D-structure.
FT DOMAIN 631 639 ASP/GLU-RICH (HIGHLY ACIDIC).
FT NP BIND 714 721 ATP (BY SIMILARITY).
FT SITE 808 811 DEAD BOX.
FT SEQUENCE 1447 AA; 163836 MW; 0DC320B41756A3C3 CRC64;

Query Match 7.6%; Score 73.5; DB 1; Length 1447;
Best Local Similarity 22.3%; Pred. No. 64;
Matches 47; Conservative 28; Mismatches 81; Indels 55; Gaps 10;

QY 1 MEPLGPEAAALPGWLALLWVSALSCPSLPASSLSLVQVQRTSYNF--GRTFLGLD 58
Db ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
934 MEPD---ERLSVQKAWQADEIQVATVAFGMGIDK-----PDVRFVYHFTVPRILEGY 985
QY 59 KCNACIGT----SICKKFF-----KEEIRSDNWL-----ASHLGLPPDLSLSYPANY 101
Db : : : : : : : : : : : : : : : : : : : : : : : : : :
986 QETGRAGRDGNYSYCITYFSFRDIRTWTQMTIQKNDRENKEHLN-KLQQVMAYCDNV 1044
QY 102 SDDSKIMRPVEIFRLVSKYQNEISDRKICASASAPKTC-----SIEVLAK 147
Db : : : : : : : : : : : : : : : : : : : : : : : : : :
1045 TDCRR-----KLVLSYFNEPDSKLC-----HKNCNCRNSANVINEERDVTPEAKK 1091
QY 148 TERPQKWLQAKRLTLPDLVQDCHQGQRELKFL 178
Db : : : : : : : : : : : : : : : : : : : : : : : : : :
1092 IVKLVESIQNERVTIIYCQDVEKGRSSKIV 1122
```

```
RESULT 13
POLR_EPMV STANDARD; PRT; 1839 AA.
AC P20126;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE RNA replicase polypotein [EC 2.7.7.48].
OS Eggplant mosaic virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Tymoviridae;
OC Tymovirus.
OX NCBI_TaxID=12151;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90021185; PubMed=2800336;
RA Osorio-Keese M.E., Keese P., Gibbs A.;
RT "Nucleotide sequence of the genome of eggplant mosaic tymovirus.";
RL Virology 172:547-554(1989).
CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC {RNA} (N).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement [See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch].
CC -----
DR EMBL; J04374; AAA43039.1; -
DR MEROPS; C21.001; -
DR InterPro; IPR008043; Peptidase C21.
DR InterPro; IPR001788; RNA_dep_RNAPol2.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR InterPro; IPR000606; Viral_helicase1.
DR Pfam; PF05381; Peptidase C21; 1.
DR Pfam; PF00978; RNA_dep_RNAPol2; 1.
DR Pfam; PF01443; Viral_helicase1; 1.
KW Transferase; RNA-directed RNA polymerase; Polyprotein; ATP-binding.
FT NP BIND 965 972 ATP (BY SIMILARITY).
FT SEQUENCE 1839 AA; 204731 MW; FD8DC1F5115E7861 CRC64;

Query Match 7.6%; Score 73.5; DB 1; Length 1839;
Best Local Similarity 20.1%; Pred. No. 86;
Matches 28; Conservative 23; Mismatches 49; Indels 39; Gaps 4;

QY 79 SDNMLASHL-----GLPPDLSLSYP-----ANYSDDS 105
Db ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
1423 SDQLLGQHLFSSLCLAYGRNPNSVLPFPQPELFSEICINDYALSSKQTATIVANHQRS 1482
QY 106 KIMRPVEIFRLVSKYQNEISDRKICASASAPKTCIER-----VLKTERFOKWLQAKRL 160
Db ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
1483 PDWRLTAV-RIFAKAOKVNDASIFSGWKACQTLALMHGYIILVLPVKKYQRIFFSKOR 1541
QY 161 TPDLVQDCHQGQRELKFLC 179
Db : : : : : : : : : : : : : : : : : : : : : : : : : :
1542 PPHIYHCGKTPSLSQMC 1560

RESULT 14
K20A HUMAN
ID K20A HUMAN STANDARD; PRT; 890 AA.
AC O95235;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Kinesin family member 20A (Rabkinesin-6) {Rab6-interacting kinesin-
DE like protein} (GG10_2).
GN KIF20A OR RAB6KIFL.
OS Homo sapiens (Human).
```


DR InterPro; IPR008957; FN_III-like.
DR InterPro; IPR003961; FN_III.
DR InterPro; IPR003962; FNIII_subd.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003598; IG_c2.
DR Pfam; PF00041; fn3; 5.
DR Pfam; PF00047; ig; 4.
DR PRINTS; PR00014; FNYPEIII.
DR SMART; SM00060; FN3; 3.
DR SMART; SM00408; IGC2; 1.
DR PROSITE; PS50835; IG-LIKE; 5.
KW Immunoglobulin domain; Muscle protein; Thick filament; Repeat.
FT DOMAIN 154 245 IG-LIKE C2-TYPE 1.
FT DOMAIN 266 371 IG-LIKE C2-TYPE 2.
FT DOMAIN 383 469 FIBRONECTIN TYPE-III 1.
FT DOMAIN 511 597 FIBRONECTIN TYPE-III 2.
FT DOMAIN 612 696 FIBRONECTIN TYPE-III 3.
FT DOMAIN 711 797 FIBRONECTIN TYPE-III 4.
FT DOMAIN 813 899 FIBRONECTIN TYPE-III 5.
FT DOMAIN 904 1002 IG-LIKE C2-TYPE 3.
FT DOMAIN 1130 1211 IG-LIKE C2-TYPE 4.
FT DOMAIN 1345 1434 IG-LIKE C2-TYPE 5.
SQ SEQUENCE 1465 AA; 164793 MW; 6F2927B5A1E69F2D CRC64;

Query Match 7.6%; Score 73; DB 1; Length 1465;
Best Local Similarity 29.0%; Pred. No. 73;
Matches 38; Conservative 15; Mismatches 50; Indels 28; Gaps 9;
QY 51 GRTFLGLDKCNACIGTS---ICKKFFKEEIRSD-----NWLASHLGL---PPD--SLLSY 97
Db 924 GNIYLGFD-CQEMTDASQFTWCKSY--EEISDDERFKIETVGDHSHKLYLKNPDKXDLGTY 980
QY 98 PANYSD-----DSKIWRPVEIFRLVSKYQNEISDRKICASAPKTCSTIERVLRKTERFQ 152
Db 981 SVSVSDTDGVSSSFVLDPEELERLMA-LSNEIKNPTI-----PLKSELAYEIFDKGRVR 1033
QY 153 KWLOAKRLTPD 163
Db 1034 FWLQAEHLSPD 1044

Search completed: June 8, 2004, 16:24:53
Job time : 19 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 8, 2004, 16:21:44 ; Search time 41 Seconds
(without alignments)
1400.593 Million cell updates/sec

Title: US-10-054-988-114
Perfect score: 962
Sequence: 1 MEPQLGPEAAALRPGWLALL.....DLVQDCHQGRELKFLCMLR 182

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL 25:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_rvirus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	962	100.0	182	4 Q9H7Y0	Q9h7y0 homo sapien
2	868	90.2	166	4 Q8WX00	Q8wx00 homo sapien
3	655	68.1	435	11 Q8C3I9	Q8c3i9 mus musculu
4	120	12.5	430	4 Q8NDZ4	Q8ndz4 homo sapien
5	91.5	9.5	384	5 Q9VWJ7	Q9vwj7 drosophila
6	86.5	9.0	438	5 O16753	O16753 caenorhabdi
7	83	8.6	398	16 Q8S646	Q8s646 escherichia
8	82.5	8.6	1107	4 Q9P2H7	Q9p2h7 homo sapien
9	81	8.4	424	16 Q8Z1G9	Q8z1g9 salmonella
10	80.5	8.4	708	12 Q88442	Q88442 strawberry
11	80	8.3	390	16 Q8XB85	Q8xb85 escherichia
12	80	8.3	390	16 Q8FA92	Q8fa92 escherichia
13	80	8.3	459	10 Q949V7	Q949v7 arabidopsis
14	80	8.3	1139	10 Q8S666	Q8s666 oryza sativ
15	80	8.3	2025	10 Q8LM71	Q8lm71 oryza sativ
16	79.5	8.3	923	16 Q833J4	Q833j4 enterococcu

17	79.5	8.3	1139	11 Q99MW8	Q99mw8 mus musculu
18	79.5	8.3	1207	11 Q9D5V1	Q9d5v1 mus musculu
19	79	8.2	424	16 Q93GQ6	Q93gq6 salmonella
20	79	8.2	551	16 Q8PBK2	Q8pbk2 xanthomonas
21	79	8.2	985	3 O59773	O59773 schizosacch
22	79	8.2	1115	10 Q9LSV7	Q9lsv7 arabidopsis
23	78.5	8.2	404	9 Q7Y442	Q7y442 enterobacte
24	78.5	8.2	504	3 Q9P3U4	Q9p3u4 schizosacch
25	78.5	8.2	1064	12 Q9YM09	Q9ym09 broad bean
26	78.5	8.2	1475	10 Q9FJV3	Q9fjv3 arabidopsis
27	78	8.1	368	12 Q90726	Q90726 human papil
28	78	8.1	398	2 O52121	O52121 escherichia
29	78	8.1	1283	10 Q8LIW4	Q8liw4 oryza sativ
30	77.5	8.1	358	13 Q7SZ55	Q7sz55 brachydanio
31	77.5	8.1	560	5 Q9U013	Q9u013 giardia lam
32	77	8.0	300	4 Q96H79	Q96h79 homo sapien
33	77	8.0	456	11 Q8VC29	Q8vcz9 mus musculu
34	77	8.0	464	11 Q9QX62	Q9qx62 mus musculu
35	77	8.0	1210	12 Q8VA41	Q8va41 chum salmon
36	77	8.0	2374	5 O45377	O45377 caenorhabdi
37	76.5	8.0	1259	10 Q9AU13	Q9aul3 oryza sativ
38	76.5	8.0	1261	10 Q8LKW0	Q8lkw0 oryza sativ
39	76	7.9	820	5 Q23215	Q23215 caenorhabdi
40	76	7.9	1006	10 Q9SRZ7	Q9srz7 arabidopsis
41	76	7.9	1697	12 Q8QZP9	Q8qzf9 crimean-con
42	76	7.9	2710	5 Q8T3G9	Q8t3g9 caenorhabdi
43	76	7.9	2712	5 O01808	O01808 caenorhabdi
44	76	7.9	7735	5 Q9VMT8	Q9vmt8 drosophila
45	75.5	7.8	260	16 Q8D8T2	Q8d8t2 vibrio vuln

ALIGNMENTS

RESULT 1

Q9H7Y0 ID Q9H7Y0 PRELIMINARY; PRT; 182 AA.
AC Q9H7Y0;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein FLJ14103.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Mammary gland;
RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,
RA Wagatsuma M., Hosoiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,
RA Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,
RA Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y.,
RA Ninomiya K., Iwayanagi T.;
RT *NEDO human cDNA sequencing project.;
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK024165; BAB14843.1; -.
KW Hypothetical protein.
SQ SEQUENCE 182 AA; 20643 MW; CA22BB5607329427 CRC64;

Query Match 100.0%; Score 962; DB 4; Length 182;
Best Local Similarity 100.0%; Pred. No. 7.6e-92;
Matches 182; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MEPQLGPEAAALRPGWLALLLWVSALSCFSLSLPASSLSLVLPQVRTSYNFGRTFLGLDKC 60

Db 1 MEPQLGPEAAALRPGWLALLLWVSALSCFSLSLPASSLSLVLPQVRTSYNFGRTFLGLDKC 60

Qy 61 NACIGTSICKFFKEEIRSDNWLASHLGLPPDLSLSPANYSDSKIWRRPVEIFRLVSKY 120

Db 61 NACIGTSICKFFKEEIRSDNWLASHLGLPPDLSLSPANYSDSKIWRRPVEIFRLVSKY 120

QY 121 QNEISDRKICASASAPKTCSTIRVLKTERFQKWLQAKRLTPDLVDCHQGORELKELCM 180
Db 121 QNEISDRKICASASAPKTCSTIRVLKTERFQKWLQAKRLTPDLVDCHQGORELKELCM 180
QY 181 LR 182
Db 181 LR 182
RESULT 2
Q8WX00
ID Q8WX00 PRELIMINARY; PRT; 166 AA.
AC Q8WX00;
DT 01-MAR-2002 (TremBLrel. 20, Created)
DT 01-MAR-2002 (TremBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TremBLrel. 20, Last annotation update)
DE BA435K1.1 (Novel protein) (Fragment).
GN BA435K1.1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Lawlor S.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL591491; CAD13520.1; -.
FT NON TER 166
SQ SEQUENCE 166 AA; 18711 MW; E23F4A20F02E74C1 CRC64;
Query Match 90.2%; Score 868; DB 4; Length 166;
Best Local Similarity 99.4%; Pred. No. 4e-82;
Matches 165; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 MEPQLGPEAAALRPGWLALLWVSALSCSFLPSSLSLVLPQVRTSYNFGRTFLGLDKC 60
Db 1 MEPQLGPEAAALRPGWLALLWVSALSCSFLPSSLSLVLPQVRTSYNFGRTFLGLDKC 60
QY 61 NACIGTSICKFFKEEIRSDNWLASHLGLPPDLSLLSYPNYSDDSKIWRPVEIFRLVSKY 120
Db 61 NACIGTSICKFFKEEIRSDNWLASHLGLPPDLSLLSYPNYSDDSKIWRPVEIFRLVSKY 120
QY 121 QNEISDRKICASASAPKTCSTIRVLKTERFQKWLQAKRLTPDLVQ 166
Db 121 QNEISDRKICASASAPKTCSTIRVLKTERFQKWLQAKRLTPDLVQ 166
RESULT 3
Q8C3I9
ID Q8C3I9 PRELIMINARY; PRT; 435 AA.
AC Q8C3I9;
DT 01-MAR-2003 (TremBLrel. 23, Created)
DT 01-MAR-2003 (TremBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TremBLrel. 24, Last annotation update)
DE Hypothetical P-loop containing nucleotide triphosphate hydrolases
DE structure containing protein.
GN 4930578C19RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
CX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Heart;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT *Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.*;
RL Nature 420:563-573 (2002).
DR EMBL; AK085770; BAC39535.1; -.
DR MGD; MGI:1923155; 4930578C19RIK.

KW Hypothetical protein.
SQ SEQUENCE 435 AA; 49042 MW; 0A1B466BB04CEB1D CRC64;
Query Match 68.1%; Score 655; DB 11; Length 435;
Best Local Similarity 78.6%; Pred. No. 1.8e-59;
Matches 132; Conservative 8; Mismatches 26; Indels 2; Gaps 2;
QY 1 MEPQ-LGPEAAALRPGWLA-ILLWVSALSCSFLPSSLSLVLPQVRTSYNFGRTFLGLD 58
Db 1 MESQWRGAAATAFHQHMLARLLWVSTLSCSFLPASLPSPVPRVRSSTYMGKTFLGLD 60
QY 59 KNCACIGTSICKFFKEEIRSDNWLASHLGLPPDLSLLSYPNYSDDSKIWRPVEIFRLVS 118
Db 61 KNCACIGTSICKFFKEEIRSDNWLASHLGLPPDLSLLSYPNYSDDSKIWRPVEISQLVS 120
QY 119 KYQNEISDRKICASASAPKTCSTIRVLKTERFQKWLQAKRLTPDLVQ 166
Db 121 RYQTEISDRRICASVSAPKTCSTIRVLKTERFQKWLQAKRLTPDLVQ 168
RESULT 4
Q8NDZ4
ID Q8NDZ4 PRELIMINARY; PRT; 430 AA.
AC Q8NDZ4;
DT 01-OCT-2002 (TremBLrel. 22, Created)
DT 01-OCT-2002 (TremBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TremBLrel. 22, Last annotation update)
DE Hypothetical protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RA Strausberg R.;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC037293; AAH37293.1; -.
KW Hypothetical protein.
SQ SEQUENCE 430 AA; 49481 MW; EB72CACF14E71457 CRC64;
Query Match 12.5%; Score 120; DB 4; Length 430;
Best Local Similarity 26.6%; Pred. No. 0.00068;
Matches 45; Conservative 29; Mismatches 77; Indels 18; Gaps 4;
QY 3 POLGPEAAALRPGWLALLWVSALSCSFLPSSLSLVLPQVRTSYNFGRTFLGLDKCNA 62
Db 7 PKLGRLSRSLKLAALGSLVLMVHSPSLIASQWQNELT-----DRFLQLNKCPA 57
QY 63 CIGTSICKFFKEEIRSDNW----LASHLGLPPDLSLLSYPNYSDDSKIWRPVEIFRLV 117
Db 58 CFGTSWCRRELNGQVVFQWGLRLDLDFLNKVN----VYFAQYGEPEGRRRVLKRLG 113
QY 118 SKYQNEISDRKICASASAPKTCSTIRVLKTERFQKWLQAKRLTPDLVQ 166
Db 114 SQRELAQLDQSICKRATGRPRCDLLQAMPRTFARLNGDVRLLTPEAVE 162
RESULT 5
Q9VWJ7
ID Q9VWJ7 PRELIMINARY; PRT; 384 AA.
AC Q9VWJ7;
DT 01-MAY-2000 (TremBLrel. 13, Created)
DT 01-MAY-2000 (TremBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TremBLrel. 24, Last annotation update)
DE CG7890 protein (RE01736p).
GN CG7890.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]

Qy 108 WRPVEIFRLVSKYQNEISDRKICASASAPKTCISIERVL-----RKTERFOK 153
Db 119 FQCSLIFREKNRKRKLIEE-----NIPKT--VEKILGIPHILFRKKERLHE 163

RESULT 7
O85646
ID O85646 PRELIMINARY; PRT; 398 AA.
AC O85646;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE L0055 (Z5142 protein) (EspG protein).
GN Z5142 OR ECS4590.
OS Escherichia coli, and
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562, 83334;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=E.coli; STRAIN=EDL933;
RX MEDLINE=98339885; PubMed=9673266;
RA Perna N.T., Mayhew G.F., Posfai G., Elliott S., Donnenberg M.S.,
RA Kaper J.B., Blattner F.R.;
RT "Molecular evolution of a pathogenicity island from enterohemorrhagic
RT escherichia coli O157:H7.";
RL Infect. Immun. 66:3810-3817(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=E.coli O157:H7; STRAIN=O157:H7 / EDL933 / ATCC 700927;
RX MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grotbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RL Nature 409:529-533(2001).
RN [3]
RP SEQUENCE FROM N.A.
RC SPECIES=E.coli O157:H7; STRAIN=O157:H7 / RIMD 0509952;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12.";
RL DNA Res. 8:11-22(2001).
DR EMBL; AF071034; AAC31534.1; -.
DR EMBL; AE005597; AAG58853.1; -.
DR EMBL; AP002566; BAB38013.1; -.
DR PIR; A86049; A86049.
DR PIR; F91202; F91202.
KW Complete proteome.
SQ SEQUENCE 398 AA; 43905 MW; F5ACE3C39DCA7DCE CRC64;

Query Match 8.6%; Score 83; DB 16; Length 398;
Best Local Similarity 27.5%; Pred. No. 4.3;
Matches 41; Conservative 19; Mismatches 51; Indels 38; Gaps 7;

Qy 7 PEAALRPGMLALLWVSALSCFSLPASSLSLVPQVRTSYNFGRTFLGLDKCNACIGT 66
Db 281 PEAT-----AICSAFYQSFNVPAITLTHERISKASEYNAERS--LDTPNACINI 328

Qy 67 SICKKFFKEEIRSDN--WLAHGLPPDLSLLSYPNYSDSKIWPRPVEIFRLVSKYQNEI 124
Db 329 SISQS-----SDGNIYVTSHTGV----LIMAPED-----RPNEMGMLTNRTSYEV 369

Qy 125 SDRKIC-----ASASAPKTCISIERVLKRT 148

Db 370 PQGVKCIIDEMVMSALQPRYAASETYLQNT 398

RESULT 8
O9P2H7
ID O9P2H7 PRELIMINARY; PRT; 1107 AA.
AC O9P2H7;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Hypothetical protein KIAA1370 (Fragment).
GN KIAA1370.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=20181126; PubMed=10718198;
RA Nagase T., Kikuno R., Ishikawa K., Hirose M., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes.XVI.
RT The complete sequences of 150 new cDNA clones from brain which code
RT for large proteins in vitro.";
RL DNA Res. 7:65-73(2000).
DR EMBL; AB037791; BAA92608.1; -.
KW Hypothetical protein.
FT NON TER 1
SQ SEQUENCE 1107 AA; 125325 MW; 393C1666F931B83B CRC64;

Query Match 8.6%; Score 82.5; DB 4; Length 1107;
Best Local Similarity 25.7%; Pred. No. 17;
Matches 39; Conservative 20; Mismatches 44; Indels 49; Gaps 9;

Qy 56 GLDK----CNACIGTSICKKFFKEEIRSDNWLASHLG-LPPDLSLLSYPNYSDSKIWPR 110
Db 667 GIDKQYSNCTT-IDKQICTNKYKEKLIENYNPKFFGNLQSD-----SKNDKIKVT 719

Qy 111 V-EIFRLVSKYQNEISDRKICASASAPKTC-----SIERVLK----- 147
Db 720 VLEMSYLNKYESMSSNK----DSKRPKTCQNTQINSIENYLNKDNKGFKCKSDQLKN 775

Qy 148 -----TERFQKWLQAKRLTPDLVQDCHQ 170
Db 776 EQDKQEDPTNEKSNYSQ-RRSIKDKLSTCEQ 806

RESULT 9
O8Z1G9
ID O8Z1G9 PRELIMINARY; PRT; 424 AA.
AC O8Z1G9;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE UV protection protein.
GN SAMB OR STY4597 OR T4291.
OS Salmonella typhi.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=601;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CT18;
RX MEDLINE=21534947; PubMed=11677608;
RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebahia M.,
RA Baker S., Basham D., Brooks K., Chillingworth T., Connerton P.,
RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
RA Feltham T., Hamlin N., Hargrett-Anderson N., Hien T.T., Holroyd S., Jagels K.,
RA Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
RA Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,
RA Whitehead S., Barrell B.G.;
RT "Complete genome sequence of a multiple drug resistant Salmonella

Db 83 NLCPLIXSSYGFKGT----DKCPYFYFSDLVVGETTCDGKKMYE-----YMAE---F 128

QY 90 PPDLSLLSYPNYSDSK--IWRPVEIFRL-----VSKYQNEIS-----DRKIC 130

Db 129 KPVHVMQLPNSVKDDASRALWK-AEMLRLQKAVEERFGHEISEDALRDAIALKNRERRAL 187

QY 131 AS-----ASAPKTCISIERVLR-----KTERF-----QKWLQAKRLTPD- 163

Db 188 ANFYHLGQFNPPALSGSDILKVVYGATFRFDKEALINELDAMTARIRQQWEEGQRLDPRP 247

QY 164 --LVQDCHQ 171

Db 248 RILITGCPIG 257

RESULT 12

Q8FA92 PRELIMINARY; PRT; 390 AA.

AC Q8FA92;

DT 01-MAR-2003 (TrEMBLrel. 23, Created)

DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)

DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)

DE Hypothetical protein yJiM.

GN YJiM OR C5418.

OS Escherichia coli O6.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

OC Enterobacteriaceae; Escherichia.

OX NCBI_TaxID=217992;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=O6:H1 / CFT073 / ATCC 700928;

RX MEDLINE=22388234; PubMed=12471157;

RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,

RA Rasko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,

RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,

RA Mobley H.L.T., Donnenberg M.S., Blattner F.R.;

RT "Extensive mosaic structure revealed by the complete genome sequence

RT of uropathogenic Escherichia coli.";

RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024 (2002).

DR EMBL; AE016772; AAN83838.1; --

KW Hypothetical protein; Complete proteome.

SQ SEQUENCE 390 AA; 43612 MW; C12BD2BFED7E83C8 CRC64;

Query Match 8.3%; Score 80; DB 16; Length 390;

Best Local Similarity 22.1%; Pred. No. 8.6;

Matches 42; Conservative 31; Mismatches 45; Indels 72; Gaps 13;

QY 39 SLVPQVRTSYNFGRTFLGLDKC-----NACIGFSIC---KKFFKEEIRSDNWLASHLGL 89

Db 83 NLCPLIXSSYGFKGT----DKCPYFYFSDLVVGETTCDGKKMYE-----YMAE---F 128

QY 90 PPDLSLLSYPNYSDSK--IWRPVEIFRL-----VSKYQNEIS-----DRKIC 130

Db 129 KPVHVMQLPNSVKDDASRALWK-AEMLRLQKAVEERFGHEISEDALRDAIALKNRERRAL 187

QY 131 AS-----ASAPKTCISIERVLR-----KTERF-----QKWLQAKRLTPD- 163

Db 188 ANFYHLGQFNPPALSGSDILKVVYGATFRFDKEALINELDAMTARIRQQWEEGQRLDPRP 247

QY 164 --LVQDCHQ 171

Db 248 RILITGCPIG 257

RESULT 13

Q949V7 PRELIMINARY; PRT; 459 AA.

AC Q949V7;

DT 01-DEC-2001 (TrEMBLrel. 19, Created)

DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE Hypothetical protein.

GN AT3G59300.

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;

OC eucots II; Brassicales; Brassicaceae; Arabidopsi.

OX NCBI_TaxID=3702;

RN [1]

RP SEQUENCE FROM N.A.

RA Yamada K., Liu S.X., Pham P.K., Banh J., Dale J.M., Goldsmith A.D.,

RA Jiang P.X., Lee J.M., Onodera C.S., Quach H.L., Tang C., Toriumi M.,

RA Yamamura Y., Yu G., Bowser L., Carninci P., Chen H., Cheuk R.,

RA Hayashizaki Y., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,

RA Kawai J., Kim C., Koesema E., Lam B., Lin J., Meyers M.C., Miranda M.,

RA Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M., Seki M.,

RA Shinn P., Southwick A., Tracy S.E., Shinozaki K., Davis R.W.,

RA Ecker J.R., Theologis A.;

RT *Full length cDNA of gene F25L23_160/AT3G59300 (GI:7801680).";

RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RA Yamada K., Banh J., Chan M.M., Chang C.H., Chang E., Dale J.M.,

RA Deng J.M., Goldsmith A.D., Lee J.M., Onodera C.S., Quach H.L.,

RA Tang C., Toriumi M., Wu H.C., Yamamura Y., Yu G., Bowser L.,

RA Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,

RA Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,

RA Meyers M.C., Miranda M., Narusaka M., Nguyen M., Palm C.J.,

RA Sakurai T., Satou M., Seki M., Shinn P., Southwick A., Shinozaki K.,

RA Davis R.W., Ecker J.R., Theologis A.;

RT *Arabidopsis Open Reading Frame (ORF) Clones.";

RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; AY050863; AAK92800.1; --

EMBL; AY091205; AAM14144.1; --

KW Hypothetical protein.

SQ SEQUENCE 459 AA; 51592 MW; B7E100293E5BF436 CRC64;

Query Match 8.3%; Score 80; DB 10; Length 459;

Best Local Similarity 31.3%; Pred. No. 10;

Matches 41; Conservative 13; Mismatches 53; Indels 24; Gaps 8;

QY 23 VSALSCSFSLPAS-----SLSSLVPQVRTSYNFGRTFLGLDKCNACIGTSICK-KPF 73

Db 2 VIAAASSFSLGRSHCHQSYTDEFSSIPYKRTSNARNRVDFGCGSANLSVLSRCKIPPF 61

QY 74 KEEIRSDNWLASH-LGLPPDLSLLSYPNYSD---DSKI--WRPVEIFRLVSKYQNEISDR 127

Db 62 GSFAHVSS--GGHDLGL---TKVSVAAADYSDSVDPSSFGYHPLLEDL----KPSKRVOET 112

QY 128 KICASASAPKT 138

Db 113 KLSASEVARTT 123

RESULT 14

Q8S666 PRELIMINARY; PRT; 1139 AA.

AC Q8S666;

DT 01-JUN-2002 (TrEMBLrel. 21, Created)

DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE Hypothetical protein.

GN OSUNBA0034B05.27.

OS Oryza sativa (japonica cultivar-group).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

OC Ehrhartoideae; Oryzeae; Oryza.

OX NCBI_TaxID=39947;

RN [1]

RP SEQUENCE FROM N.A.

RA Wing R.A., Yu Y., Soderlund C., Chen M., Kim H.-R., Rambo T.,

RA Sasaki C., Henry D., Oates R., Simmons J.;

RT *Rice Genomic Sequence.";

RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

Search completed: June 8, 2004, 16:25:48
Job time : 43 secs

SEQUENCE 1139 AA; 127148 MW; 7F50B31552E999FD CRC64;

Matches 37; Conservative 18; Mismatches 50; Indels 34; Gaps 7;

Db 336 PEFAAVESGPL-ILAWAVFLCLVMSLPGESNTNLGPVSGFRGILRTTFISAFVASYEISYQT 394

Db 395 EDSSLCMILNLCBVYDGEESLCMOFWDKDSFIDGPIRFNYLERMNGV--TTLVAVPRSD 452

Db 453 TDNVNYHDOIEIHSPISIF 471

RA Wing R.A., Yu Y., Yang T.J., Nah G., Soderlund C., Chen M., Kim H.-R.,
 RA Rambo T., Saski C., Henry D., Oates R., Simmons J.;
 RT "Rice Genomic Sequence.";
 RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.

Best local similarity 20.08, P-Seq. no. 03,
Matches 37; Conservative 18; Mismatches 50;
Indels 34; Gaps 7;

db 335 PEEAAVESGPI-I LAWAVFLCLVMSLPGSNTNLGPVSGFERGIIRTFISAFVASYEISYOT 394

הב 395 ENSSJ.GMT*NTI.CEVVNGEEST.CMOENDKDSEFTNGPTBENVI.ERMNEY--TTI.YAVPRSH 452

453 TONNAGE: 2375 TUCIDCTE 471